

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

ANALYSE DE L'ARGUMENTATION DE LA VALIDITÉ DES INFÉRENCES
D'ÉVALUATION DANS LES POLITIQUES INSTITUTIONNELLES
D'ÉVALUATION DES APPRENTISSAGES DES ÉTABLISSEMENTS
D'ENSEIGNEMENT COLLÉGIAL QUÉBÉCOIS

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN ÉDUCATION

PAR
KARINE PAQUETTE-CÔTÉ

MARS 2010

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 -Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

À Gilles Raïche, le directeur espéré. Pour votre considération, votre pertinence, votre connaissance, votre éthique, votre confiance, votre respect, merci !

À Frédéric Legault, pour avoir fait des embûches un chemin pavé, merci. À mes évaluateurs, pour le temps accordé, les possibles casse-têtes et pour vos conseils avisés, merci. À Gérald Boutin, pour avoir alimenté et encouragé mon inclination pour la méthodologie et la recherche, merci.

Pour les ébats d'esprit, et les plus érudits, pour les accolades et les rigolades, pour l'apéro ou le tarot, merci à Danielle, André-Jacques, Pierre, Violaine, Chantal, François, Bastien, Hélène, Michel, Bernard, Martin.

Pour votre présence malgré les longues absences, pour votre intérêt, vos attentions, votre support, votre compréhension, pour m'avoir fait reconnaître ma fierté à chaque étape achevée, et pour tout le plaisir et l'amour que vous pouvez me faire éprouver, merci à Sylvia, Steve, Marco, Myriam, Jon, François, Mylène, David, Pascale, Pascal, Julie, Yannick, Justine, P-O, Mathieu, Sonia, Christian, Vlad, Alain, France.

Avec vous, il est toujours 9h05.

Louise et Paul, quand on dit « les mots me manquent », parce qu'avec vous aucun mot n'est jamais nécessaire, puisqu'ils sont toujours ressentis... Considérez ceci comme votre accomplissement au même titre que le mien.

À Danielle Paquette et Serge Côté, vous êtes les plus importants artisans de qui je suis et de ce que j'accomplis.

TABLE DES MATIÈRES

LISTE DES FIGURES.....	vii
LISTE DES TABLEAUX.....	viii
RÉSUMÉ	ix
INTRODUCTION	1
CHAPITRE I	
CONTEXTE THÉORIQUE.....	6
1.1 L'assurance de qualité de l'évaluation des apprentissages de la Commission d'évaluation de l'enseignement collégial : les politiques institutionnelles d'évaluation des apprentissages (PIEA).....	7
1.2 L'assurance de qualité de l'évaluation des apprentissages au niveau théorique : la notion de validité.....	11
1.2.1 La validation de l'argument interprétatif proposée par Kane (2006).....	13
1.2.1.1 Le trait	14
1.2.1.2 Le domaine cible	15
1.2.1.3 L'univers de généralisation	16
1.2.1.4 La validation de l'argument interprétatif pour l'interprétation d'un trait.....	19
1.3 La validité de l'évaluation des apprentissages	25
1.3.1 Les pratiques d'évaluation des apprentissages dans le réseau collégial québécois	25

1.3.2 La validation de l'évaluation des apprentissages selon l'approche par compétences	27
1.3.2.1 Une validation <i>a priori</i> par recours à des juges.....	29
1.3.2.2 Une validation empirique interne	30
1.3.2.3 Une validation empirique externe	31
1.3.2.4 La validation et la preuve de la maîtrise de la compétence	32
1.4 L'assurance de qualité de l'évaluation des apprentissages dans les établissements d'enseignement collégial québécois à travers l'analyse de la structure d'argumentation dans les politiques institutionnelles d'évaluation des apprentissages (PIEA).....	33
1.4.1 Objectif de recherche.....	35
CHAPITRE II MÉTHODOLOGIE.....	36
2.1 Corpus d'analyse.....	36
2.2 Instrumentation	38
2.2.1 Grille d'analyse	38
2.2.2 Outil de modélisation schématique	39
2.3 Déroulement.....	40
2.3.1 Analyse de contenu	40
2.3.2 Modélisation schématique.....	41
2.4 Méthode d'analyse et d'interprétation des données	43
2.4.1 Une analyse quantitative de statistiques descriptives.....	43
2.4.2 Une analyse qualitative de synthèse	43
2.5 Considérations éthiques	44
CHAPITRE III RÉSULTATS	45
3.1 Description et répartition des énoncés	46

3.2 Validation de la structure d'argumentation interprétative de Kane (2006).....	49
3.2.1 Exhaustivité des catégories du modèle de Kane (2006).....	49
3.2.2 Exclusivité des catégories du modèle de Kane (2006).....	52
3.2.3 Pertinence des catégories du modèle de Kane (2006).....	53
3.2.4 Synthèse des résultats du premier objectif	54
3.3 Identification de moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages	57
3.3.1 Inférence de notation (I1)	57
3.3.1.1 Les règles de passation et de notation sont appropriées (A1.1)	58
3.3.1.2 Les règles de passation et de notation sont appliquées telles que spécifiées (A1.2).....	60
3.3.1.3 La passation et la notation sont exemptes de biais (A1.3)	62
3.3.1.4 Les données s'adaptent à tous les modèles d'échelles de mesure employés dans la notation (A1.4).....	65
3.3.2 Inférence de généralisation (I2).....	66
3.3.2.1 L'échantillon des observations est représentatif de l'univers de généralisation (A2.1)	67
3.3.2.2 L'échantillon des observations est suffisamment grand pour contrôler l'erreur aléatoire (A2.2).....	70
3.3.3 Inférence d'extrapolation (I3)	71
3.3.3.1 Le score univers est lié au score cible (A3.1).....	72
3.3.3.2 Il n'y a pas d'erreur systématique qui soit susceptible de miner l'extrapolation (A3.2).....	75
3.3.4 Inférence d'implication (I4)	77
3.3.4.1 Les implications associées au trait sont appropriées (A4.1)....	78
3.3.4.2 Les propriétés des scores observés supportent les implications associées avec l'étiquette du trait (A4.2)	79
3.3.5 Crédibilité de l'inférence d'évaluation (C)	82
3.3.5.1 L'ensemble des arguments est respecté de façon à ce que chacune des inférences considérées de façon individuelle soit convaincante (A5.1)	83

3.3.5.2 Les processus d'apprentissage et d'évaluation sont connus des acteurs (A5.2)	84
3.3.6 Répartition des moyens selon les niveaux d'inférence d'évaluation	88
CHAPITRE IV	
DISCUSSION ET INTERPRÉTATIONS	90
4.1 Les lacunes au niveau de la mesure	90
4.2 Le problème de l'intégration de compétences auxiliaires à l'évaluation	95
4.3 Le problème de l'évaluation des dispenses, des substitutions et des équivalences	99
4.4 La nature des moyens identifiés	101
CONCLUSION	104
RÉFÉRENCES	108
APPENDICE A	
GRILLE D'ANALYSE DU CONTENU MANIFESTE DES POLITIQUES INSTITUTIONNELLES D'ÉVALUATION DES APPRENTISSAGES	113
APPENDICE B	
MOYENS POUR CHERCHER À ASSURER LA VALIDITÉ DES INFÉRENCES D'ÉVALUATION DES APPRENTISSAGES	116

LISTE DES FIGURES

Figure	Page
1.1 Procédure de mesure et d'argumentation interprétative pour l'interprétation d'un trait (adapté de Kane, 2006, p. 33).....	13
1.2 Les quatre niveaux d'inférence de l'interprétation d'un trait et les arguments garantissant leur validité.	22
2.1 Représentation graphique de l'argument A1.2 <i>Les règles de passation et de notation sont appliquées telles que spécifiées</i>	42
3.1 Répartition des énoncés en fréquences et nombre de moyens identifiés selon les catégories du modèle d'argumentation de la validité des inférences d'évaluation	47
3.2 Structure d'argumentation de la validité des inférences d'évaluation des apprentissages	56
3.3 Inférence de notation (I1).....	58
3.4 Inférence de généralisation (I2)	66
3.5 Inférence d'extrapolation (I3)	71
3.6 Inférence d'implication (I4)	77
3.7 Crédibilité de l'inférence d'évaluation.....	82

LISTE DES TABLEAUX

Tableau		Page
1.1	L'argument interprétatif pour l'interprétation d'un trait (adapté de Kane , 2006, p. 34)	20
2.1	Description des établissements d'enseignement collégial dont les politiques institutionnelles d'évaluation des apprentissages constituent le corpus d'analyse	36
3.1	Réduction des données à travers le processus de catégorisation	50
3.2	Répartition des énoncés selon le nombre de catégories dans lesquels ils ont été classés.....	52

RÉSUMÉ¹

Cette recherche vise, d'un point de vue scientifique, à valider la structure d'argumentation interprétative de Kane (2006) par l'application de cette structure à l'analyse de politiques institutionnelles d'évaluation des apprentissages (PIEA) du réseau collégial québécois. Elle vise aussi, d'un point de vue pratique, à identifier des moyens que peuvent mettre en place les institutions pour permettre d'assurer l'argumentation de la validité des inférences au regard des apprentissages des étudiants. Une analyse de contenu de politiques institutionnelles d'évaluation des apprentissages (PIEA) est réalisée à partir de la structure d'argumentation interprétative de Kane (2006). Une modélisation schématique des moyens identifiés dans les politiques institutionnelles d'évaluation des apprentissages (PIEA) est produite, menant à l'élaboration de lignes directrices permettant de chercher à assurer l'argumentation de la validité des inférences d'évaluation en évaluation des apprentissages au collégial. Cette recherche a aussi permis de générer des hypothèses quant à l'exhaustivité, l'exclusivité et la pertinence des catégories du modèle de Kane (2006) dans ce contexte, tout en apportant une suggestion en complément à ce modèle théorique.

MOTS-CLÉS : structure d'argumentation interprétative, validité, inférence d'évaluation, évaluation des apprentissages, politiques institutionnelles d'évaluation des apprentissages (PIEA)

¹ Les normes de présentation utilisées dans ce mémoire correspondent à l'adaptation francophone des normes APA rédigée par Gilles Raïche en 2007 pour la *Revue des sciences de l'éducation*. Selon ces normes de présentation, l'introduction, plus élaborée, tient aussi lieu de problématique.

INTRODUCTION

À l'enseignement supérieur, les apprentissages devraient être authentiques et significatifs puisque ceux-ci sont en lien direct avec les actes et la profession future ou actuelle des étudiants. Les modalités d'évaluation des apprentissages doivent donc se rapprocher des pratiques du milieu professionnel et des situations réelles auxquelles les étudiants, professionnels et futurs professionnels, devront faire face. (Gipps, 1994 ; Scallon, 2004 ; Wiggins, 1993)

En 1993, la notion de compétence s'est répandue à l'ordre collégial dans le système d'éducation du Québec en référence à la nécessité d'apprentissages authentiques et significatifs. Selon Raïche (2008), la compétence s'apparente à un objectif d'apprentissage, tel qu'on en a traditionnellement traité en évaluation des apprentissages. Toutefois, le terme compétence est plus précis et réfère à « *un enseignement significatif et authentique, un apprentissage significatif et authentique ainsi qu'une évaluation significative et authentique de ce dernier* » (Raïche, 2008, p. 2). L'application de l'enseignement par compétences s'est traduite par diverses mesures, dont la détermination des objectifs et des standards des programmes

d'études en fonction des compétences par le ministère de l'Éducation², la mise en place d'épreuves synthèse pour les programmes d'étude et la mise en œuvre d'actions de concertations entre les ordres d'enseignement.

Au niveau de l'évaluation, qu'il s'agisse de l'évaluation des apprentissages traditionnels ou de l'évaluation des compétences, il est aujourd'hui reconnu que le processus d'évaluation est toujours limité quant à son potentiel d'authenticité face à la tâche réelle (Raïche, 2008). Le processus d'évaluation demeure une simulation et la performance de l'étudiant à celle-ci se distingue nécessairement de sa performance en contexte réel. À ce propos, Muchielli (1971, p. 30-31) rappelle qu'un test est une mise à l'épreuve expérimentale qui correspond à une situation simulée. Lorsqu'on réfère à la correspondance entre la performance en contexte réel et le résultat obtenu à un test, on réfère principalement à la notion de validité.

Traditionnellement, la validité en évaluation a souvent été définie comme la capacité d'un test à bien mesurer ce qu'il prétend mesurer. En 1951, Cureton définissait la validité comme ayant deux dimensions, la pertinence et la fiabilité (*relevance and reliability*), et comme étant la corrélation entre le score obtenu au test et le « vrai » score critérié (*"true" criterion score*) (p. 623). En 1966, l'American Psychological Association propose trois types de validité : la validité de construit, la validité de

² Le ministère de l'Éducation porte le nom de ministère de l'Éducation, du Loisir et du Sport depuis le remaniement ministériel du 18 février 2005 (<http://www.mels.gouv.qc.ca>).

contenu et la validité critériée, cette dernière étant prédictive ou concomitante (Cronbach, 1971). Cette conception de la validité ayant été développée pour la mesure en psychologie, Cronbach (1971) souligne le besoin de définir la validité en fonction de son utilisation et de son interprétation en éducation. La validité est encore principalement associée à la validité d'un test, mais plus d'importance est accordée à l'interprétation. Cronbach (1971, p. 447) mentionne d'ailleurs que parce que chaque interprétation a son propre degré de validité, on ne peut jamais arriver à la simple conclusion qu'un test particulier « est valide »³. Il ajoute ensuite que tous les aspects et tous les détails d'une procédure de mesure peuvent influencer la performance et donc ce qui est mesuré (p. 449). Il place la validité dans un contexte de validation, laquelle correspond au processus visant à évaluer la précision des prédictions ou des inférences réalisées sur la base du résultat obtenu à un test. En 1989, Messick définit la validité comme jugement évaluatif intégré du degré auquel la preuve empirique et le rationnel théorique supportent la justesse et la pertinence des inférences et des actions basées sur les résultats d'une évaluation⁴. Cette définition implique une conception de la validité en tant qu'argument, telle que l'a introduite Cronbach (1980) qui décrit la validation du jugement évaluatif comme un processus rhétorique dans lequel l'évaluateur doit justifier son jugement par la présentation d'arguments réalistes fondés sur des preuves empiriques. Deux principaux modèles ont été

³ « Because every interpretation has its own degree of validity, one can never reach the simple conclusion that a particular test "is valid." » (Cronbach, 1971, p. 447)

⁴ « Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and actions based on test scores or other modes of assessment » (Messick, 1989, p. 13).

développés en parallèle dans les années 1990 et 2000 en accord avec cette conception de la validité : l'un centré davantage sur la preuve empirique et l'autre davantage sur le rationnel théorique. Le premier est une approche de conception de l'évaluation basée sur la preuve, le modèle ECD (« *evidence-centred design* ») développée par Mislevy et ses collaborateurs (Mislevy, Almond et Lukas, 2004 ; Mislevy et Haertel, 2006a, 2006b ; Mislevy, Steinberg, Almond et Lukas, 2006). Le second est une structure d'argumentation interprétative développée par Kane (1992 ; 2004 ; 2006), qui peut aussi être considérée comme un modèle méthodologique, pouvant servir à la validation des arguments interprétatifs en évaluation des apprentissages. Le modèle de Kane et celui de Mislevy et de ses collaborateurs, ainsi que celui de Toulmin (1964) et les travaux de Messick (1989) qui les ont inspirés, sont d'ailleurs aujourd'hui des références incontournables relativement au concept de la validité en mesure et évaluation (Lissitz, 2009). Toutefois, la mise à l'épreuve empirique de ces modèles théoriques reste encore à faire. Au moment de la rédaction du présent mémoire de recherche, les études portant sur les applications de ces modèles récents ne font que débiter.

En somme, il est souhaitable qu'à l'enseignement supérieur l'apprentissage soit significatif et authentique (Gipps, 1994 ; Scallon, 2004 ; Wiggins, 1993) et qu'une évaluation significative et authentique soit faite de ce dernier (Raïche, 2008). Les modèles développés autour de la notion de validité en tant que jugement évaluatif intégré nécessitent une mise à l'épreuve empirique afin de pouvoir contribuer

concrètement à la qualité de l'enseignement et de l'évaluation en éducation. Il semble donc légitime de poser la question suivante : *comment peut-on assurer la validité de l'interprétation des résultats de l'évaluation des apprentissages des étudiants à l'enseignement supérieur ?*

La présente recherche tente de répondre à cette question à travers l'exploration des politiques institutionnelles d'évaluation des apprentissages (PIEA) qui orientent les pratiques en évaluation des apprentissages en enseignement supérieur, et plus spécifiquement en enseignement collégial. C'est ce qu'expose le Chapitre I de ce mémoire, ainsi que certains développements théoriques concernant la validité en mesure et évaluation en éducation. De cet exposé, découlent deux objectifs de recherche qui sont formulés en fin de chapitre. L'atteinte de ces objectifs a été réalisée au moyen d'une méthodologie qualitative décrite au Chapitre II. Les résultats issus d'analyses quantitatives descriptives et qualitatives de synthèse sont présentés au Chapitre III, dont certains sont discutés au Chapitre IV.

Cette recherche de nature exploratoire a permis l'élaboration d'hypothèses et de suggestions théoriques, ainsi que d'outils pratiques pour contribuer à assurer la validité de l'interprétation des résultats de l'évaluation des apprentissages des étudiants à l'enseignement supérieur.

CHAPITRE I

CONTEXTE THÉORIQUE

Le développement des pratiques et des théories en éducation au cours des dernières décennies présente un intérêt croissant envers la qualité de l'évaluation des apprentissages. Le présent chapitre expose certaines tentatives visant l'atteinte de cette dimension; d'abord par un survol du développement des instances et des politiques d'évaluation dans le réseau d'enseignement collégial du Québec, ensuite à travers l'exposé du développement de la notion de validité en évaluation des apprentissages et d'un modèle théorique, celui de Kane (2006), permettant d'assurer l'argumentation de la validité des inférences d'évaluation des apprentissages. Un regard est ensuite posé sur les résultats d'une étude réalisée par Howe et Ménard (1993), portant sur les croyances et les pratiques en évaluation des apprentissages dans le réseau collégial québécois, dans le but d'estimer si les pratiques actuelles au sein du réseau d'enseignement collégial permettent la validation des inférences d'évaluation. Puis, pour vérifier que le modèle de validation de l'argument interprétatif de Kane (2006) est compatible avec les pratiques d'évaluation de l'approche par compétences, actuellement en place dans le réseau collégial québécois, les recommandations de Ketele et Gerard (2005) concernant la validation des épreuves d'évaluation selon l'approche par compétences sont étudiées. Ce chapitre se termine par la formulation des objectifs de la présente recherche.

1.1 L'assurance de qualité de l'évaluation des apprentissages de la Commission d'évaluation de l'enseignement collégial : les politiques institutionnelles d'évaluation des apprentissages (PIEA)

En 1993 avait lieu le renouveau du réseau collégial du Québec. La grande visée du renouveau était d'« *assurer aux jeunes et à l'ensemble de la population du Québec un enseignement collégial d'un calibre et d'une qualité qui leur permettent de se mesurer aux meilleurs standards de compétence* » (Ministère de l'Enseignement supérieur et de la Science, 1993, p. 13). Le renouveau introduisait dans le système scolaire québécois la notion de compétence en référence à la nécessité d'apprentissages authentiques et significatifs. Le renouveau, tout en accordant davantage d'autonomie aux institutions, a donné lieu à la mise en place de mesures visant à garder un certain contrôle sur la qualité de la formation et de l'évaluation au sein du réseau collégial.

Avec le renouveau en 1993 fût sanctionnée la loi 83 et avec elle fût instituée la Commission d'évaluation de l'enseignement collégial (CÉEC) qui est un organisme autonome et indépendant relevant directement de l'Assemblée nationale du Québec. La Commission d'évaluation de l'enseignement collégial fût créée pour contribuer au développement de la qualité, de la crédibilité et de la reconnaissance de la formation offerte dans les établissements d'enseignement collégial québécois. La Commission d'évaluation de l'enseignement collégial devait resserrer les mécanismes d'évaluation des pratiques institutionnelles tant au niveau de l'évaluation des apprentissages que

de l'évaluation des programmes d'études. Dès l'automne 1994, chaque établissement d'enseignement collégial du Québec était dans l'obligation d'adopter les mesures suivantes (Commission d'évaluation de l'enseignement collégial, 1994a, p. 4) :

- *définir et appliquer une politique institutionnelle d'évaluation des apprentissages (PIEA);*
- *définir et appliquer une politique institutionnelle d'évaluation des programmes d'études (PIEP);*
- *soumettre à la Commission, pour des fins d'évaluation, sa politique institutionnelle d'évaluation des apprentissages (PIEA) et sa politique institutionnelle d'évaluation des programmes d'études (PIEP);*
- *collaborer avec la Commission à l'évaluation de la mise en œuvre des programmes d'étude qu'il dispense.*

Dès lors, la Commission d'évaluation de l'enseignement collégial est chargée « *de porter un jugement formel de qualité sur la manière dont les collèges remplissent leurs responsabilités académiques* » (Ministère de l'Enseignement supérieur et de la Science, 1993, p. 27).

Dans le domaine de l'évaluation des apprentissages, ce sont les politiques institutionnelles d'évaluation des apprentissages (PIEA) qui permettent à la Commission d'évaluation de l'enseignement collégial de porter un jugement sur la qualité de l'évaluation des apprentissages des étudiants dans les établissements d'enseignement collégial. Une politique institutionnelle d'évaluation des apprentissages (PIEA) est :

Un instrument qui permet à un collège d'orienter, d'encadrer et de soutenir les activités reliées à l'évaluation des apprentissages. Elle établit les objectifs poursuivis par le collège, les principes et les valeurs qui orientent les actions, ainsi que les responsabilités de tous les groupes concernés. (Conseil des collèges, 1992, p. 38)

L'évaluation des politiques institutionnelles d'évaluation des apprentissages (PIEA) par la Commission d'évaluation de l'enseignement collégial consiste à « *évaluer pour attester la qualité de l'évaluation des apprentissages et pour contribuer à l'améliorer* » (Commission d'évaluation de l'enseignement collégial, 1994b, p. 3). Ainsi, la Commission d'évaluation de l'enseignement collégial s'assure que les objectifs d'apprentissage et les seuils de réussite sont clairement définis, qu'ils peuvent être évalués et que les pratiques d'évaluation sont pertinentes, cohérentes, efficaces et transparentes. La Commission d'évaluation de l'enseignement collégial doit aussi s'assurer que l'évaluation des apprentissages est équitable et donc s'assurer de l'équivalence de la formation et des diplômes. La Commission d'évaluation de l'enseignement collégial (1994b, p. 7) considère comme équitable l'évaluation qui est « *fidèle au contenu enseigné* » et qui est « *équivalente pour un même cours dispensé par des enseignants et des enseignantes différents* ». L'évaluation laisse une marge d'autonomie aux enseignants, mais elle est sous la responsabilité institutionnelle, c'est-à-dire que l'institution est responsable de soutenir et d'appuyer les enseignants dans l'exercice de leurs fonctions d'évaluateurs. Les moyens que doivent mettre en place les institutions à cette fin doivent être spécifiés par les institutions dans les politiques institutionnelles d'évaluation des apprentissages (PIEA).

Chaque établissement d'enseignement collégial doit donc rédiger et appliquer sa propre politique institutionnelle d'évaluation des apprentissages (PIEA), permettre à la Commission d'évaluation de l'enseignement collégial de procéder à leur évaluation et appliquer les modifications recommandées par cette dernière. Selon la Commission d'évaluation de l'enseignement collégial (1994b, p. 11) une politique institutionnelle d'évaluation des apprentissages (PIEA) doit contenir les composantes essentielles suivantes : les finalités et les objectifs, les moyens, le partage des responsabilités, et enfin, les moyens et les critères de l'auto-évaluation de l'application de la politique. Cette dernière composante réfère aux processus et aux actions prévus par l'établissement d'enseignement collégial pour procéder à l'évaluation de sa propre politique institutionnelle d'évaluation des apprentissages (PIEA). La Commission d'évaluation de l'enseignement collégial (1994b, p. 11) suggère que cette auto-évaluation soit réalisée à partir des critères suivants : « *la conformité de l'application avec le texte de la politique, l'efficacité de cette application pour garantir la qualité de l'évaluation des apprentissages et l'équivalence de l'évaluation des apprentissages pour contribuer à en assurer l'équité* ». Suivant le deuxième critère mentionné, les établissements doivent veiller à ce que leur politique institutionnelle d'évaluation des apprentissages (PIEA) permette de garantir la qualité de l'évaluation des apprentissages. Les écrits théoriques du domaine de l'évaluation des apprentissages fournissent des moyens de répondre à ce critère, notamment en ce qui concerne la notion de validité.

1.2 L'assurance de qualité de l'évaluation des apprentissages au niveau théorique : la notion de validité

Comme il a été mentionné plus tôt, le processus d'évaluation est limité quand à son potentiel d'authenticité face à la tâche réelle (Raïche, 2008). La mesure en évaluation implique l'utilisation d'échantillons limités d'observations pour dresser des conclusions générales et abstraites sur la ou les compétences d'un individu dans un domaine plus ou moins défini. Ces conclusions ou interprétations tirées de l'évaluation sont décisives pour la formation et la certification des étudiants se préparant au milieu professionnel; elles doivent par conséquent être plausibles et appropriées.

La plausibilité et la justesse des conclusions tirées de l'évaluation des apprentissages font référence à la notion de validité, ou plus précisément à la validation de l'interprétation de l'évaluation. La fonction première de l'évaluation est de porter un jugement et ce jugement doit être plausible et cohérent avec les inférences issues de la procédure de mesure (Cronbach, 1971). Ces inférences sont produites à partir d'évidences elles-mêmes produites à partir d'une procédure de mesure. Cronbach, Linn, Brennan et Haertel (1997, p. 376) rappellent que la mesure consiste à faire l'évaluation d'un échantillon de performance. Ce qui nous intéresse dans l'évaluation en éducation est d'évaluer le niveau de compétence d'un étudiant dans un domaine de compétence et non seulement dans un échantillon restreint de performances. Ils

précisent que le domaine de généralisation, c'est-à-dire jusqu'à quel point et à quel ensemble d'observations les résultats à la mesure peuvent être généralisés, dépend des utilisations et des interprétations qui seront faites de ces résultats.

Les évidences ou preuves empiriques ne peuvent à elles seules constituer une validation du jugement évaluatif. Comme il a été mentionné dans l'introduction, pour Messick (1989), la validité est un jugement évaluatif intégré du degré auquel la preuve empirique et le rationnel théorique supportent la justesse et la pertinence des inférences et des actions basées sur les résultats d'une évaluation⁵. Les évidences ou preuves empiriques doivent être incorporées dans un argument logique (Cronbach, 1980). Cronbach décrit la validation du jugement évaluatif comme un processus rhétorique dans lequel l'évaluateur doit justifier son jugement par la présentation d'arguments réalistes fondés sur des preuves empiriques. En 1988, il propose de baser la validation de l'interprétation et de l'utilisation des résultats des évaluations sur la logique de l'argument évaluatif (Cronbach, 1988). Se basant sur cette proposition et sur la structure de l'argumentation élaborée par Toulmin (1964), Kane (1992) introduit la notion d'argumentation interprétative à travers une approche de validation fondée sur l'argumentation. En 2006, il propose une structure d'argumentation interprétative, qui peut aussi être considérée comme un modèle méthodologique,

5 « Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and actions based on test scores or other modes of assessment » (Messick, 1989, p. 13).

pouvant servir à la validation des arguments interprétatifs en évaluation des apprentissages.

1.2.1 La validation de l'argument interprétatif proposée par Kane (2006)

La figure 1.1 présente la structure d'argumentation interprétative pour l'interprétation d'un trait développée par Kane (2006). Le texte qui suit est destiné à la description des éléments constitutifs de cette figure.

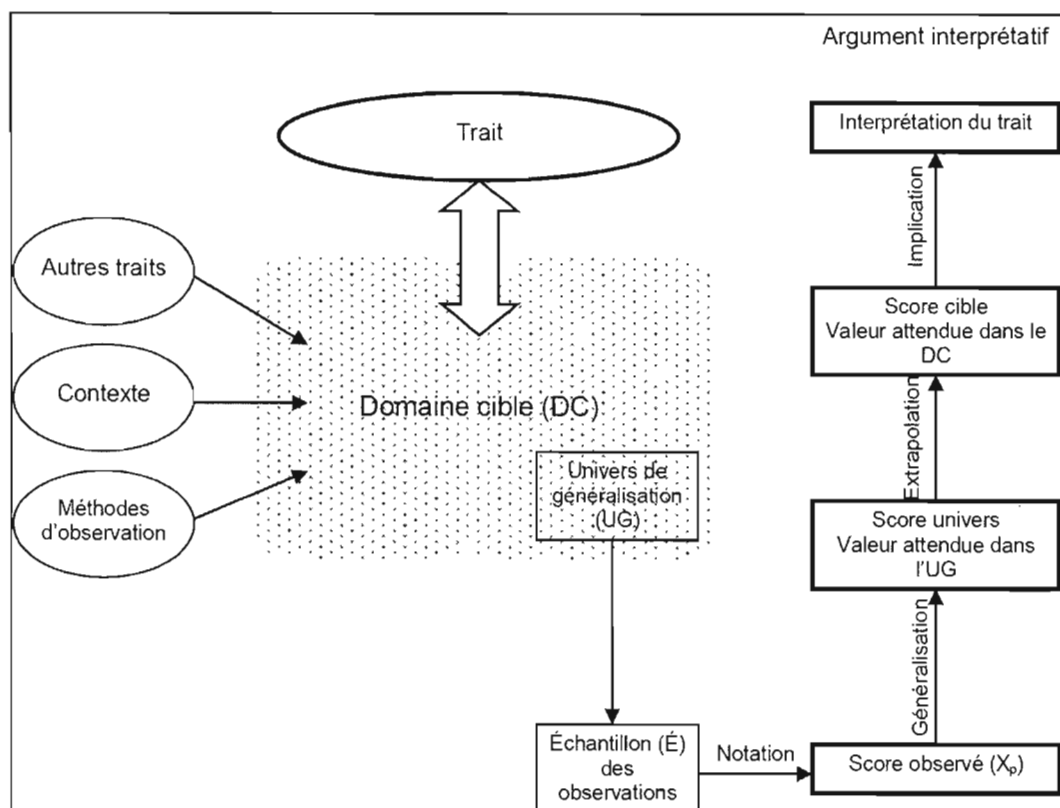


Figure 1.1 Procédure de mesure et d'argumentation interprétative pour l'interprétation d'un trait (adapté de Kane, 2006, p. 33)

1.2.1.1 Le trait

Selon Kane, un *trait* est une disposition à agir ou à performer d'une certaine façon en réponse à certains types de stimuli ou de tâches, sous certaines circonstances. La signification d'un trait est donnée par le domaine à partir duquel il est défini, mais l'interprétation d'un trait implique aussi, implicitement du moins, que certains attributs sous-jacents ou latents expliquent les régularités observées dans la performance (Loevinger, 1957 : voir Kane, 2006, p. 30). On peut s'attendre à ce que les personnes qui présentent un trait fort performent bien à des tâches ou des situations liées à ce trait. Kane (2006) considère le trait comme un attribut unidimensionnel. Il mentionne que dans la plupart des cas où une terminologie du trait est utilisée, l'attribut mesuré implique une combinaison de composantes pratiques, mais les fonctions de l'attribut sont considérées comme s'il n'y avait qu'un seul trait unidimensionnel qui opèrerait de la même façon en tant que déterminant du succès pour tous les items d'un test (Henrysson, 1971, p. 146 : voir Kane, 2006, p. 32).

On peut rapprocher le terme trait au terme compétence qui est plus couramment utilisé en évaluation des apprentissages dans les établissements d'enseignement québécois. En effet, la notion de compétence peut elle aussi être considérée comme unidimensionnelle tout en impliquant une combinaison de composantes pratiques et elle implique elle aussi la régularité de la performance « *dans une famille de*

situations comparables qui exigent leur intégration » (Scallon, 2004, p. 204). De plus, la notion de compétence est en lien direct avec les actes et la profession future ou actuelle des étudiants (Raïche, 2008) ; sa signification est donc donnée par le domaine auquel elle est associée, tout comme c'est le cas pour le trait.

1.2.1.2 Le domaine cible

Selon les termes de Kane (2006), le *domaine cible* est un ensemble d'observations liées au trait. Il s'agit de l'ensemble des manifestations du trait, dans les divers contextes dans lesquels le trait est impliqué, ainsi que des liens qui l'unissent à d'autres traits et d'autres domaines. La décision d'inclure certaines observations, plutôt que d'autres, dans un domaine cible est généralement liée à l'expérience ou à des suppositions sur les processus impliqués dans les observations. Ainsi, par exemple, l'analyse d'un article scientifique peut être une tâche incluse dans le domaine cible de la méthodologie de la recherche parce qu'on présume qu'elle requiert les mêmes habiletés ou des habiletés connexes ou des performances constitutives (*component performances*). Le fait que cela puisse être difficile d'inclure certaines tâches dans une évaluation ne signifie pas qu'elles ne font pas partie du domaine cible d'un trait. L'objectif n'est pas de définir un domaine bien ordonné ou un domaine qui soit facile à évaluer, mais d'identifier la gamme des observations associées avec la compétence ou le trait d'intérêt. Le *score cible*

correspond à la valeur attendue d'une personne dans le domaine cible, c'est-à-dire à la performance attendue d'une personne à l'ensemble des observations pouvant être associées au trait d'intérêt. Le score cible résulte donc de l'interprétation du score obtenu par une personne en lien avec le domaine cible. En évaluation des apprentissages, le score cible serait donc une estimation de la compétence de l'étudiant dans le domaine, résultant de l'interprétation du ou des scores obtenus par un étudiant aux tests ou aux mesures. Le score en lui-même n'est pas différent, mais son interprétation est étendue au domaine. La correspondance entre le domaine cible et la mesure du trait est un sujet central dans le développement de la procédure d'évaluation et dans la validation de l'interprétation des résultats de l'évaluation.

1.2.1.3 L'univers de généralisation

Dans la plupart des cas en évaluation des apprentissages, il est quasi impossible de former un échantillon aléatoire ou représentatif des observations à partir du domaine cible et de généraliser le *score observé* d'une personne dans cet échantillon (ex. : résultat obtenu à un test) au score attendu d'une personne dans le domaine cible (*score cible*). La gamme des observations incluses dans l'évaluation est par conséquent plus restreinte que celle du domaine cible. Conséquemment, les observations incluses dans les mesures des traits sont typiquement tirées d'un sous-ensemble du domaine cible, souvent un très petit sous-ensemble (Fitzpatrick et

Morrison, 1971 : voir Kane, 2006; Kane, 1982 : voir Kane, 2006). Kane (2006) nomme ce sous-ensemble l'*univers de généralisation*, qui est un échantillon des observations du domaine cible à partir duquel sont tirées les observations qui constitueront la mesure (par exemple, les items d'un examen, les critères d'évaluation d'un stage, etc.). En évaluation des apprentissages, l'univers de généralisation représente donc le programme d'étude ou de façon plus restreinte le cours. Le score attendu d'une personne à travers l'univers de généralisation est le *score univers*. C'est l'interprétation du score en lien avec l'univers de généralisation, donc, en évaluation des apprentissages, l'estimation de la compétence de l'étudiant dans le cours ou dans le programme d'étude. Alors que le domaine cible pour la compétence d'une personne en méthodologie de la recherche devrait inclure, par exemple, l'examen critique et l'application de la méthodologie dans une grande variété de matériel écrit (ex. : journaux d'actualité, manuels d'études, magazines, articles de revues scientifiques, rapports de recherche), de réponses (ex. : répondre à des questions spécifiques, produire un résumé écrit ou oral, expérimenter une méthode, produire un rapport de recherche), et de contextes (ex. : en classe, à la maison, à la bibliothèque, en milieu professionnel), l'univers de généralisation pour une mesure de la compétence en méthodologie de la recherche peut parfois être limité aux réponses à des questions objectives portant sur la lecture d'un article scientifique en contexte d'examen en classe ou à l'application d'une méthode dans la réalisation d'un travail de session. Ces performances constituent effectivement des exemples d'observations liées à la méthodologie de la recherche, mais elles forment une parcelle très restreinte

du domaine cible de la compétence en méthodologie de la recherche. Le score observé, dans ce cas-ci le résultat obtenu à l'examen ou au travail de session, peut dans de telles circonstances bien représenter le score univers, la performance attendue de la personne en contexte d'examen en classe ou de réalisation d'un travail de session. Toutefois, on peut se demander s'il reflète avec justesse le score cible, soit la performance attendue de la personne en méthodologie de la recherche; sachant que le domaine cible, la méthodologie de la recherche, inclut une variété beaucoup plus vaste d'observations que celles incluses dans l'univers de généralisation. C'est ce questionnement, la représentativité du score observé par rapport au score cible, qui est à la base de la validation de l'argument interprétatif pour l'interprétation d'un trait.

1.2.1.4 La validation de l'argument interprétatif pour l'interprétation d'un trait

L'interprétation de la performance observée en termes de score cible requiert une chaîne de raisonnements de la performance à la mesure au score observé, du score observé au score univers, et du score univers au score cible. De plus, l'interprétation d'un trait porte avec elle son lot d'implications, lesquelles peuvent inclure des relations avec d'autres variables, l'impact des interventions sur ce trait, les exceptions à l'interprétation et l'étendue ou l'importance des différences entre les groupes (Cronbach, 1971, p. 448 : Kane, 2006, p. 32). La plupart des interprétations ont des implications qui vont au-delà du domaine cible et qui nécessitent d'être validées. Le tableau 1.1, décrit de façon plus détaillée par la suite, présente une vue d'ensemble de l'argument interprétatif pour l'interprétation d'un trait tel que décrit par Kane (2006).

Tableau 1.1 L'argument interprétatif pour l'interprétation d'un trait (adapté de Kane, 2006, p. 34)

I1 : Notation (<i>scoring</i>) : à partir de la performance à la mesure au score observé
A1.1 Les règles de passation et de notation sont appropriées.
A1.2 Les règles de passation et de notation sont appliquées telles que spécifiées.
A1.3 La passation et la notation sont exemptes de biais.
A1.4 Les données s'adaptent à tous les modèles d'échelles de mesure employés dans la notation.
I2 : Généralisation : à partir du score observé au score univers
A2.1 L'échantillon des observations est représentatif de l'univers de généralisation.
A2.2 L'échantillon des observations est suffisamment grand pour contrôler l'erreur aléatoire.
I3 : Extrapolation : à partir du score univers au score cible
A3.1 Le score univers est lié au score cible.
A3.2 Il n'y a pas d'erreur systématique qui soit susceptible de miner l'extrapolation.
I4 : Implication : à partir du score cible à la description verbale de l'interprétation du trait
A4.1 Les implications associées au trait sont appropriées.
A4.2 Les propriétés des scores observés supportent les implications associées avec l'étiquette du trait.

L'argument interprétatif représenté à droite de la figure 1.1 comprend quatre niveaux d'inférence pour l'interprétation des résultats d'un test en termes de trait : la notation, la généralisation, l'extrapolation et l'implication (I1 à I4 dans le tableau 1.1). Lors de l'inférence de notation (I1), la performance observée dans l'échantillon des observations qui constituent la mesure est notée en se rapportant à un score observé (un score brut ou un score sur une échelle). Par l'inférence de généralisation (I2), le score observé est généralisé au score univers, c'est-à-dire que le score obtenu à la

mesure est interprété dans un contexte plus large qui est l'univers de généralisation. Par l'inférence d'extrapolation (I3), le score univers est extrapolé au score cible, c'est-à-dire que le score est interprété non plus comme représentatif de la performance à un sous-ensemble d'observations, mais comme étant représentatif du domaine en entier. Finalement, par l'inférence d'implication (I4), les implications associées au trait sont jointes à l'interprétation du score dans le domaine de façon à décrire ses relations avec d'autres traits et d'autres domaines, à faire certaines mises en garde quant aux exceptions ou aux diverses interprétations pouvant être produites au regard du trait en question ou du score. Ces quatre niveaux d'inférence (I1 à I4 dans le tableau 1.1) sont accompagnés des arguments (A1.1 à A4.2 dans le tableau 1.1) qui garantissent la validité de chaque inférence et, par le fait même, la validité de l'interprétation du trait. Les quatre niveaux d'inférence accompagnés de leurs arguments de validité sont représentés à la figure 1.2 et expliqués ensuite en détails, de l'inférence de notation à l'inférence d'implication.

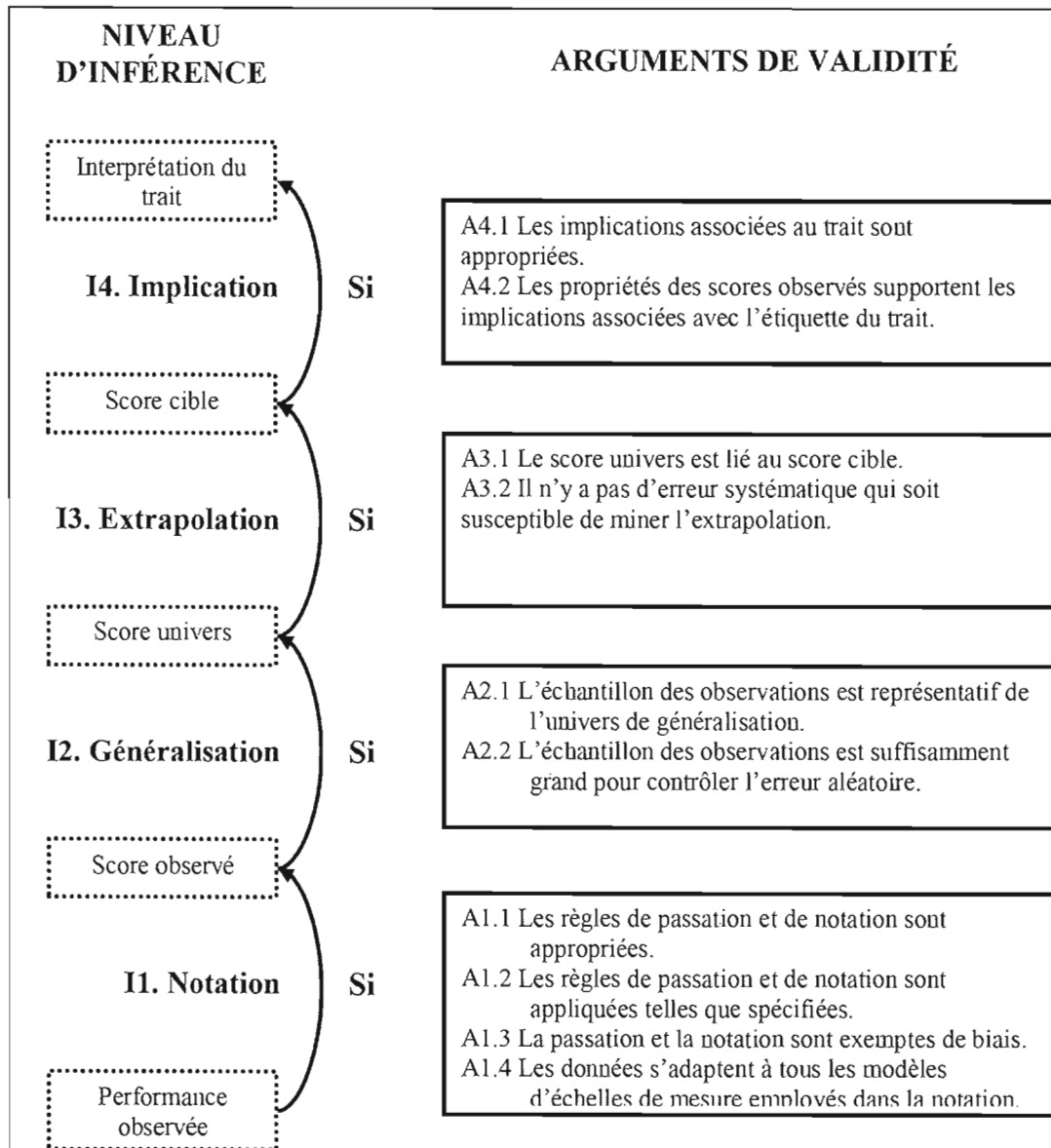


Figure 1.2 Les quatre niveaux d'inférence de l'interprétation d'un trait et les arguments garantissant leur validité.

L'inférence de notation assigne un score à la performance de chaque personne en utilisant une règle de notation, laquelle fournit la garantie (*warrant* en anglais) de la validité l'inférence de notation. On assume alors que les critères de notation sont

appropriés et qu'ils sont appliqués tels que prévu, que le processus de passation et de notation est exempt de biais, et que tous les modèles statistiques (échelle de mesure, calcul des scores) employés dans la notation sont appropriés.

L'inférence de généralisation étend l'interprétation du score observé à partir de l'évaluation d'un échantillon d'observations, qui constitue la mesure, à la valeur attendue à travers l'univers de généralisation, soit le score univers. La valeur du score est la même, mais son interprétation s'étend d'une affirmation sur un ensemble spécifique d'observations à une affirmation sur la performance attendue à travers l'univers de généralisation. La garantie de cette inférence est dérivée de la théorie de l'échantillonnage statistique et dépend du fait que l'on assume la représentativité de l'échantillon des observations et la convenance de la taille de l'échantillon pour contrôler l'erreur d'échantillonnage. Ainsi, pour que la généralisation de l'interprétation du score observé au score univers soit considérée comme valide, on doit s'assurer que les items sélectionnés pour constituer la mesure sont en nombre suffisant pour permettre de contrôler l'erreur aléatoire et qu'ils représentent bien l'ensemble des items constituant l'univers de généralisation.

L'extrapolation à partir de l'univers de généralisation au domaine cible étend l'interprétation du score univers au score cible. Dans l'inférence d'extrapolation, on assume que le score univers est lié (rationnellement et/ou empiriquement) au score cible et que l'extrapolation est relativement exempte d'erreurs systématiques ou

d'erreurs aléatoires. Encore une fois, le score ne change pas, mais l'interprétation du score s'étend de l'univers de généralisation au domaine cible.

Les implications regroupent les extrapolations de l'interprétation pour y inclure toutes affirmations ou suggestions associées au trait. Les garanties des implications du trait fournissent l'autorisation pour toute inférence pouvant être formulée ou impliquée dans la description du trait, dans l'étiquette du trait, et dans les utilisations faites des résultats de l'évaluation. Par exemple, en adoptant une étiquette de trait déjà existante, un terme du langage courant associé à un sens communément accepté, le développeur du test adopte implicitement dans l'interprétation proposée la signification étendue du terme ou alors il se voit obligé de s'opposer à toute inférence non garantie pouvant être faite sur la base de l'étiquette du trait. Ou encore, s'il est attendu par exemple que le trait demeure stable dans le temps, les résultats empiriques qui laissent croire qu'il en est effectivement ainsi devraient supporter l'interprétation; dans le cas contraire, l'implication devrait spécifier que la stabilité du trait dans le temps n'a pas été démontrée et qu'elle n'est pas garantie.

Pour que l'argument interprétatif décrit au tableau 1.1 et représenté à la figure 1.2 soit convaincant, chacune des inférences considérées de façon individuelle doit être convaincante. Une faille dans n'importe laquelle des inférences produites invalide l'argument interprétatif en entier, et ce, même si les évidences pour les autres inférences ont bien été amenées (Crooks et al., 1996 : voir Kane, 2006, p. 34).

1.3 La validité de l'évaluation des apprentissages

1.3.1 Les pratiques d'évaluation des apprentissages dans le réseau collégial québécois

Une étude menée par Howe et Ménard (1993), dans le cadre du Programme d'aide à la recherche sur l'enseignement et l'apprentissage (PAREA) du ministère de l'Éducation, fait état des croyances et des pratiques en évaluation des apprentissages dans le réseau collégial québécois. Leur étude permet de constater que certaines pratiques en évaluation des apprentissages ne facilitent pas l'interprétation de la performance observée en termes de score cible. Autrement dit, il est loin d'être évident que la performance des étudiants telle que mesurée dans les classes des établissements d'enseignement collégial soit représentative de leur compétence dans le domaine cible. Les indices que fournissent Howe et Ménard (1993) quant à ce doute sont directement liés au non respect des arguments de validité de Kane (2006), qui invalident l'entière interprétation du trait. Bien que l'étude de Howe et Ménard (1993) ait été menée avant le renouveau du réseau collégial de 1993 et la mise en œuvre de l'approche par compétences, la nature des pratiques observées laissent supposer que celles-ci pourraient ne pas avoir changé de manière substantielle et qu'elle serait donc encore appliquées de nos jours dans les établissements d'enseignement collégial.

Howe et Ménard (1993) mentionnent entre autres que la majorité des enseignants du collégial intègrent à leur mesure des cibles d'évaluation qui ne sont pas liées au trait, aux apprentissages ou à la compétence ciblée. Ces cibles d'évaluation, telles que le respect des exigences de présentation des travaux, la motivation et l'effort à l'apprentissage, la participation active de l'étudiant en classe et la présence assidue aux cours, sont plutôt tel que l'affirment Howe et Ménard (1994) des stratégies de contrôle des étudiants qui ne sont pas reliées aux apprentissages ciblés :

En incorporant dans les notes des composantes autres que les résultats d'apprentissage, on diminue la validité de l'information consignée au bulletin de l'étudiant et on rend cette information difficile à interpréter. (p. 24)

Ainsi, selon les termes de Kane (2006), l'échantillon des observations utilisé dans la mesure d'évaluation n'étant pas représentatif de l'univers de généralisation (voir A2.1 dans le tableau 1.1), l'inférence de généralisation du score observé au score univers est invalide, ce qui invalide par le fait même l'entière interprétation du tait.

Howe et Ménard (1994, p. 25) mentionnent aussi que, malgré le fait que 90 % des enseignants croient que tous les étudiants d'un même cours devraient être évalués sur la base des mêmes critères, « *plusieurs enseignants ont tendance à modifier la note finale d'un étudiant, soit pour le pénaliser, mais plus souvent pour le récompenser.* » En modifiant ainsi la note de l'étudiant, l'inférence de notation est biaisée puisque les règles de notation ne sont pas appliquées telles que spécifiées (voir A1.2 dans le

tableau 1.1), ce qui, comme le mentionnent Howe et Ménard (1994, p. 26), « *rend difficile l'interprétation des notes consignées au bulletin de l'élève.* »

Comme le démontrent les résultats de Howe et Ménard (1993), certaines pratiques d'évaluation des apprentissages dans le réseau collégial ne permettent pas la validation de l'argument interprétatif pour l'interprétation d'un trait parce qu'elles présentent des failles d'inférence.

À la lecture de ces constats, il serait légitime de se demander si le modèle de validation de l'argument interprétatif pour l'interprétation d'un trait de Kane (2006) demeure compatible avec les pratiques d'évaluation de l'approche par compétences, approche actuellement en place dans le réseau collégial québécois.

1.3.2 La validation de l'évaluation des apprentissages selon l'approche par compétences

De Ketele et Gerard (2005) ont soulevé certains problèmes relatifs à l'utilisation des techniques classiques de validation des évaluations dans l'approche par compétences. Un des problèmes identifiés est l'utilisation psychométrique traditionnelle d'échantillons d'items représentatifs lors de la construction de la mesure. Comme il a déjà été mentionné, la mesure en évaluation implique l'utilisation d'échantillons

limités d'observations pour dresser des conclusions générales et abstraites sur la ou les compétences d'un individu dans un domaine plus ou moins défini. De Ketele et Gerard (2005) soulignent que les évaluations élaborées selon l'approche par compétences risquent de présenter des problèmes de validité et de fidélité en raison de la taille limitée de l'échantillon issu de l'univers de référence. Selon ces auteurs, l'approche par compétence ne rend pas possible la validation des outils d'évaluation selon les critères de validités issus de la psychométrie, dont les suivants :

- *le caractère représentatif de l'échantillon d'items par rapport à l'univers de référence ;*
- *la position de la moyenne (pas trop éloignée du centre de l'échelle de notation) ;*
- *la dispersion des résultats (variance la plus grande possible pour discriminer au mieux les élèves) ;*
- *la forme de la distribution (courbe gaussienne) ;*
- *l'unidimensionnalité du trait mesuré (en l'occurrence, une aptitude scolaire spécifique). (De Ketele et Gerard, 2005, p. 7)*

Ils affirment que les épreuves utilisées dans l'approche par compétences sont pertinentes parce qu'elles proposent des situations complexes plus significatives et authentiques, plus près de la réalité, mais qu'« *il est difficile d'en attester la validité parce que les méthodes pour assurer cette validation ne sont pas évidentes ou tout simplement pas connues.* » (De Ketele et Gerard, 2005, p. 5)

De Ketele et Gerard (2005, p. 9) proposent alors l'usage des types de validations suivantes :

- *une validation a priori, par recours à des juges ;*
- *une validation empirique interne ;*
- *une validation empirique externe.*

Ces types de validations ne sont pas incompatibles avec le modèle de validation des arguments interprétatifs en évaluation des apprentissages proposé par Kane (2006).

1.3.2.1 Une validation *a priori* par recours à des juges

De Ketele et Gerard (2005) proposent de procéder à la validation *a priori* des items de l'évaluation en demandant à des juges de vérifier si les items ou les situations-problèmes de l'évaluation font bien partie de l'univers de référence. Pour ce faire, ils suggèrent de disposer d'une description précise de l'univers de référence, c'est-à-dire d'une définition des paramètres qui caractérisent la famille des situations-problèmes. Le rapprochement des notions de trait et de compétence a déjà été abordé (voir 1.2.1.1 Le trait, p. 10). Il a été mentionné que, selon Kane (2006), la signification d'un trait est donnée par le domaine à partir duquel il est défini et que la décision d'inclure certaines observations, plutôt que d'autres, dans un domaine cible est généralement liée à l'expérience ou à des suppositions sur les processus impliqués dans les observations. La validation *a priori* par le recours à des juges est donc une

technique tout à fait compatible avec le modèle de validation des arguments interprétatifs proposé par Kane (2006) puisque des juges, experts du domaine cible, possèdent effectivement l'expérience nécessaire pour juger des items ou des situations-problèmes qui font effectivement partie du domaine cible. Ainsi, dans une approche par compétences, l'argument de validité « *A2.1 L'échantillon des observations est représentatif de l'univers de généralisation.* » de Kane (2006) pourrait être formulé ainsi : « *L'échantillon des paramètres de la situation est représentatif des paramètres de la famille des situations* ». Dans un cas comme dans l'autre, il demeure possible pour des juges de décider des observations ou des paramètres de la famille de situations qui constituent l'univers de généralisation.

1.3.2.2 Une validation empirique interne

L'approche par compétences ne permet pas une validation interne traditionnelle par analyse factorielle ou par calcul de coefficients d'homogénéité parce que l'évaluation n'est pas constituée d'un ensemble dénombré d'items correspondant à une dimension unique. Pour procéder à la validation interne d'une évaluation par compétences, De Ketele et Gerard (2005) proposent de fonder la validation sur les critères d'évaluation qui sont invariants aux situations d'une même famille. Ainsi, si un étudiant obtient un score élevé à un critère d'évaluation dans une situation d'évaluation, il devrait obtenir un score élevé au même critère dans une autre situation d'évaluation de même

famille. À cette validation empirique interne peuvent être associés tous les arguments de validité de l'inférence de notation (voir I1 dans le tableau 1.1) du modèle de Kane (2006). Il s'agit en effet de s'assurer que les critères ou les règles de passation et de notation sont appropriés, qu'ils sont appliqués tels que spécifiés, qu'ils sont exempts de biais et que les données recueillies s'adaptent aux échelles de mesure employés dans la notation. En ce qui concerne la validation empirique interne, le modèle de validation des arguments interprétatifs en évaluation des apprentissages proposé par Kane (2006) ne semble pas incompatible avec l'approche par compétences.

1.3.2.3 Une validation empirique externe

De Ketele et Gerard (2005) mentionnent qu'il est possible d'effectuer une validation empirique externe en comparant la performance d'un même groupe à une épreuve compétence et à une évaluation portant sur une ou plusieurs variables critères pour lesquelles des relations ont été démontrées avec la compétence évaluée. En comparant la performance des étudiants à une épreuve compétence, qui est un échantillon tiré de l'univers de généralisation, à leur performance à une évaluation d'une variable critère reliée à la compétence, qui est un autre échantillon tiré de l'univers de généralisation, on s'assure, lorsque les deux performances sont corrélées, que les deux types d'évaluations mesurent bien une même réalité. Le score observé

dans l'épreuve peut dans de telles circonstances bien représenter le score univers, mais on ne peut s'assurer qu'il représente bien le score cible que si le lien entre le score univers et le score cible a été démontré et qu'il n'y a pas eu d'erreurs systématiques ou d'erreurs aléatoires (voir A3.1 et A3.2 dans le tableau 1.1).

1.3.2.4 La validation et la preuve de la maîtrise de la compétence

Selon De Ketele et Gerard (2005), la maîtrise de chacun des critères d'une épreuve bien construite ne certifie pas la maîtrise de la compétence visée. Il s'agit là d'une préoccupation à la base de la démarche de validation de l'inférence d'évaluation. En effet, le modèle de Kane (2006) propose une démarche méthodologique permettant de s'assurer de la validité de chacune des inférences passant par la notation, la généralisation, l'extrapolation et l'implication du trait mesuré. Cette démarche s'apparente au concept de validité globale élaboré par Auger et Ségin (1996) et définit par Auger (2003) ainsi :

Évaluation de la contribution de l'ensemble des validités spécifiques, de l'opérationnalisation des critères de scientificité, des protocoles mis en place, fondée sur des évidences empiriques et sur un rationnel théorique, en vue d'assurer l'adéquation et la justesse des inférences et des actions à partir d'informations recueillies ou de scores au test.

C'est la démonstration de la validité de chacune des inférences qui rend valide l'argument interprétatif en entier, c'est-à-dire qui permet d'établir avec rigueur le lien entre le score observé et le trait ciblé.

1.4 L'assurance de qualité de l'évaluation des apprentissages dans les établissements d'enseignement collégial québécois à travers l'analyse de la structure d'argumentation dans les politiques institutionnelles d'évaluation des apprentissages (PIEA).

L'étude de Howe et Ménard (1993) démontre que certaines pratiques d'évaluation des apprentissages dans le réseau collégial ne permettent pas la validation de l'argument interprétatif pour l'interprétation d'un trait parce qu'elles présentent des failles d'inférence. La structure d'argumentation décrite par Kane (2006) devrait théoriquement permettre de soutenir la validation des inférences d'évaluation. Comme une politique institutionnelle d'évaluation des apprentissages (PIEA) est un outil devant permettre d'orienter, d'encadrer et de soutenir les pratiques d'évaluation des apprentissages dans les établissements d'enseignement collégial, il semble justifié de s'attendre à ce qu'elles contiennent les éléments permettant la validation des inférences d'évaluation. Elles devraient donc contenir les éléments présentés dans le modèle de Kane (2006) tels que, par exemple, des spécifications sur les règles de notation appropriées, sur la réduction des biais associés à la notation, sur la

représentativité des observations, etc. L'étude des recommandations de de Ketele et Gerard (2005) concernant la validation des épreuves d'évaluation selon l'approche par compétences démontre que le modèle de validation des arguments interprétatifs en évaluation des apprentissages proposé par Kane (2006) ne semble pas incompatible avec l'approche par compétences qui est en place actuellement dans le réseau collégial québécois. De plus, l'application du modèle de Kane (2006) pourrait éventuellement permettre de répondre à une des préoccupations de de Ketele et Gérard (2005), soit l'établissement du lien entre le score obtenu à l'évaluation et le trait ou la compétence ciblée. L'analyse des politiques institutionnelle d'évaluation des apprentissages (PIEA) selon la structure d'argumentation décrite par Kane (2006) pourrait être un pas de plus pour assurer la qualité de l'évaluation des apprentissages des étudiants dans nos établissements d'enseignement.

1.4.1 Objectif de recherche

Le rapprochement des politiques institutionnelles d'évaluation des apprentissages (PIEA) et du modèle théorique de Kane (2006), deux types d'assurances de qualité de l'évaluation des apprentissages, l'une pratique et l'autre théorique, mène à la formulation de l'objectif principal de ce projet qui est de valider la structure d'argumentation interprétative de Kane (2006) par l'application de cette structure à l'analyse de politiques institutionnelles d'évaluation des apprentissages (PIEA) du réseau collégial québécois. Un objectif secondaire de cette recherche est d'identifier, à la lueur du contenu des politiques analysées, des moyens que peuvent mettre en place les institutions pour chercher à assurer la validité des inférences d'évaluation au regard des apprentissages des étudiants.

CHAPITRE II

MÉTHODOLOGIE

2.1 Corpus d'analyse

Le corpus d'analyse de la présente recherche est constitué de cinq politiques institutionnelles d'évaluation des apprentissages (PIEA). Le tableau 2.1 présente la description des établissements d'enseignement collégial dont la politique institutionnelle d'évaluation des apprentissages (PIEA) a été analysée.

Tableau 2.1 Description des établissements d'enseignement collégial dont les politiques institutionnelles d'évaluation des apprentissages (PIEA) constituent le corpus d'analyse

PIEA	Population de la municipalité (en 2006) ^a	Public / Privé	Francophone / Anglophone	Nombre d'étudiants (en 2006) ^b
01	Moins de 50 000	Public	Francophone	Environ 2500
02	Moins de 50 000	Public	Francophone	Moins de 1500
03	Plus de 1 000 000	Public	Francophone	Environ 2500
04	Plus de 1 000 000	Privé	Francophone	Moins de 1500
05	Plus de 1 000 000	Public	Anglophone	Plus de 5000

^a Données issues de Statistiques Canada (<http://www.statcan.gc.ca>)

^b Données issues des *Statistiques détaillées sur l'éducation* (Ministère de l'Éducation, des Loisirs et du Sport, 2007).

Pour des raisons d'homogénéité, seules les politiques institutionnelles d'évaluation des apprentissages (PIEA) s'appliquant à la formation ordinaire, par opposition à la formation continue, ont été considérées. Les collèges privés non subventionnés ne sont pas pris en compte par les analyses, car leur contexte éducationnel diffère des collèges subventionnés en ce qui a trait à leur nombre d'étudiants généralement inférieur et à leur système de gestion et d'organisation particulier. Sur les 64 collèges privés et publics, anglophones et francophones, offrant la formation ordinaire et figurant sur la liste fournie par le Ministère de l'Éducation, des Loisirs et du Sport, 41 rendent disponible leur politique institutionnelle d'évaluation des apprentissages (PIEA) sur leur site Internet. Les cinq politiques institutionnelles d'évaluation des apprentissages (PIEA) qui constituent le corpus d'analyse ont été sélectionnées à partir de ces dernières et de façon à inclure des collèges privés et publics, anglophones et francophones, de moyennes et grandes populations de la province de Québec.

Les résultats de la présente recherche sont limités aux établissements d'enseignement collégial dont la politique institutionnelle d'évaluation des apprentissages (PIEA) a été analysée. Toutefois, puisque la Commission d'évaluation de l'enseignement collégial met à la disposition de la population et des collèges un cadre de référence sur l'évaluation des politiques institutionnelles d'évaluation des apprentissages (PIEA) (Commission d'évaluation de l'enseignement collégial, 1994), le contenu des politiques institutionnelles d'évaluation des apprentissages (PIEA) des

établissements d'enseignement collégial québécois est relativement homogène. De ce fait, il est attendu que les interprétations issues des résultats de la présente recherche puissent s'appliquer à des établissements d'enseignement collégial dont la politique institutionnelle d'évaluation des apprentissages (PIEA) n'a pas été analysée.

2.2 Instrumentation

L'instrumentation de la présente recherche est constituée d'une grille d'analyse de contenu et d'un outil de modélisation schématique.

2.2.1 Grille d'analyse

Une grille d'analyse de contenu (voir l'appendice A) a été construite à partir de la structure argumentaire développée par Kane (2006) et présentée au tableau 1.1 qui permet d'assurer la validité de l'argumentation interprétative pour l'interprétation d'un trait. Cette grille d'analyse est constituée de cinq colonnes. La première colonne permet d'identifier la politique institutionnelle d'évaluation des apprentissages (PIEA) analysée. La deuxième colonne contient l'unité d'analyse, soit un énoncé correspondant à une phrase ou un ensemble de phrases exprimant une seule idée. Dans la troisième colonne est identifié pour chaque extrait l'argument de validité du modèle de Kane (2006) auquel il fait référence. Dans cette colonne, l'unité

d'analyse peut être codée par un seul ou par plusieurs arguments de validité du modèle de Kane (2006). L'unité d'analyse peut aussi être codée « Autre » lorsque l'énoncé se rapporte à la validité de l'inférence d'évaluation, mais qu'il ne correspond à aucun argument du modèle de Kane (2006) en particulier. La quatrième colonne permet de coder si un moyen d'application et/ou d'évaluation de l'argument est identifié dans l'énoncé. La cinquième colonne est réservée aux interprétations de l'analyste pouvant être, selon le cas et sans s'y limiter : le moyen d'application et/ou d'évaluation de l'argument, une interprétation au sujet de l'énoncé ou de l'argument de validité auquel il fait référence, une interprétation au sujet du modèle de Kane (2006).

2.2.2 Outil de modélisation schématique

Une fois que chaque énoncé a été codé dans la grille d'analyse, les moyens d'application et/ou d'évaluation identifiés dans les politiques institutionnelles d'évaluation des apprentissages (PIEA) analysées sont modélisés à l'aide du logiciel MotPlus développé par le Centre de recherche LICEF de la Télé-université (LICEF). MotPlus est un éditeur graphique qui « *vise à soutenir une méthode de représentation graphique des connaissances développée à l'intention des concepteurs pédagogiques.* » (Rivard, 2007).

2.3 Déroulement

La méthodologie utilisée dans la réalisation de ce mémoire est centrée sur l'analyse de contenu des politiques institutionnelles d'évaluation des apprentissages (PIEA). Une première étape consiste à analyser le contenu manifeste des politiques institutionnelles d'évaluation des apprentissages (PIEA) au moyen d'une grille d'analyse. Une deuxième étape de l'analyse consiste à modéliser les informations recueillies au moyen de la grille d'analyse.

2.3.1 Analyse de contenu

La première étape de l'analyse consiste à analyser le contenu manifeste des politiques institutionnelles d'évaluation des apprentissages (PIEA). Pour les fins de l'analyse de contenu, les textes sont découpés en énoncés distincts, chaque énoncé devenant une unité d'analyse. Un énoncé correspond à une phrase ou un ensemble de phrases exprimant une seule idée. À cette étape, ne sont conservés que les éléments les plus significatifs contenus dans les textes étudiés (Lessard-Hébert, Goyette et Boutin, 1996; Van der Maren, 1995). Le processus de catégorisation ou de classification suit un *modèle fermé* (L'Écuyer, 1987), c'est-à-dire que les catégories sont prédéterminées, issues de la structure argumentaire permettant la validation de l'interprétation d'un trait décrite par Kane (2006).

2.3.2 Modélisation schématique

Un processus de modélisation est ensuite réalisé permettant de classer au moyen d'une représentation schématique les moyens pouvant être mis en place par les acteurs des établissements d'enseignement collégial pour chercher à assurer la validité des inférences d'évaluation au regard des apprentissages des étudiants.

À l'étape de la modélisation, le processus de catégorisation correspond à un *modèle mixte* (L'Écuyer, 1987, p. 59) permettant d'effectuer des regroupements, des subdivisions ou des ajouts au modèle de Kane (2006). La structure de base correspond à la représentation du modèle de Kane présentée à la figure 1.2. Chaque argument de validité (A1.1 à A4.2) est développé en sous-modèle consistant en une représentation graphique de l'ensemble des moyens qui lui sont associés. Chaque moyen identifié à l'étape de l'analyse de contenu est ajouté à la représentation graphique de l'argument auquel il correspond. Les moyens d'application et/ou d'évaluation sont classés et présentés en fonction des acteurs responsables de leur application et/ou de leur évaluation. La figure 2.1 illustre à titre d'exemple la représentation graphique réalisée pour l'argument *A1.2 Les règles de passation et de notation sont appliquées telles que spécifiées*.

2.4 Méthode d'analyse et d'interprétation des données

Deux analyses sont réalisées à partir des données recueillies : une analyse quantitative de statistiques descriptives et une analyse qualitative de synthèse.

2.4.1 Une analyse quantitative de statistiques descriptives

Une analyse quantitative de statistiques descriptives (fréquences absolues et relatives, moyennes, écarts types) est effectuée à partir des données obtenues pour évaluer l'exhaustivité, l'exclusivité et la pertinence des catégories du modèle original de Kane (2006) dans le contexte de l'évaluation des apprentissages en enseignement collégial, permettant ainsi de répondre à l'objectif principal de la présente recherche. Cet objectif est de valider la structure d'argumentation interprétative de Kane (2006) par l'application de cette structure à l'analyse de politiques institutionnelles d'évaluation des apprentissages (PIEA) du réseau collégial québécois.

2.4.2 Une analyse qualitative de synthèse

Le second objectif de cette recherche est d'identifier des moyens que peuvent mettre en place les institutions pour chercher à assurer la validité des inférences d'évaluation au regard des apprentissages des étudiants. Une analyse qualitative de synthèse est

réalisée pour faire ressortir les résultats les plus significatifs pour chaque argument de la structure d'argumentation de la validité des inférences d'évaluation au regard des moyens relevés lors de l'analyse de contenu. L'analyse a été décrite plus en en détails dans les sections *2.3.1 Analyse de contenu* (p. 40) et *2.3.2 Modélisation schématique* (p. 41). Les moyens conservés sont formulés sur la base du contenu des politiques institutionnelles d'évaluation des apprentissages (PIEA) et de façon à représenter tous les niveaux de la modélisation schématique.

2.5 Considérations éthiques

L'analyse des données de la recherche respecte l'anonymat des établissements au regard des interprétations. Elle ne vise pas l'évaluation et le classement des établissements.

Les résultats de cette recherche sont communiqués au réseau collégial et à la communauté de la recherche en éducation. Ce mémoire de recherche sera transmis à la Fédération des cégeps ainsi qu'à la Commission d'évaluation de l'enseignement collégial. Il a déjà fait l'objet de trois communications en congrès (Paquette-Côté, 2009 ; Paquette-Côté et Raïche, 2008 ; Paquette-Côté et Raïche, 2009) et d'un article dans un ouvrage collectif sur les mécanismes pour assurer la validité de l'interprétation de la mesure en éducation (Paquette-Côté et Raïche, à paraître).

CHAPITRE III

RÉSULTATS

Une analyse de contenu de cinq politiques institutionnelles d'évaluation des apprentissages (PIEA) a été réalisée et le présent chapitre expose les résultats de cette analyse. La section 3.1 présente la description et la répartition globale des énoncés analysés. La section 3.2 présente les résultats de l'analyse quantitative de statistiques descriptives pour répondre à l'objectif principal de la présente recherche qui porte sur la validation la structure d'argumentation interprétative de Kane (2006). Celle-ci est centrée sur l'évaluation de l'exhaustivité, l'exclusivité et la pertinence des catégories du modèle original de Kane (2006) dans le contexte de l'évaluation des apprentissages en enseignement collégial. Cette section se termine par une synthèse des résultats liés au premier objectif de la présente recherche. La section 3.3 présente les résultats de l'analyse qualitative de synthèse faisant ressortir les résultats les plus significatifs pour chaque argument de la structure d'argumentation de la validité des inférences d'évaluation au regard des moyens relevés lors de l'analyse de contenu. Cette section se termine par la description de la répartition des moyens identifiés à travers les divers niveaux d'inférence de la structure d'argumentation de la validité des inférences d'évaluation.

3.1 Description et répartition des énoncés

Un total de 424 énoncés tirés des politiques institutionnelles d'évaluation des apprentissages (PIEA) ont été analysés. La figure 3.1, expliquée plus en détails par la suite, présente le nombre d'énoncés classés dans chaque catégorie pour chaque politique institutionnelle d'évaluation des apprentissages (PIEA) ainsi que les moyennes et écarts types. Rappelons qu'un même énoncé pouvait être classé dans plus d'une catégorie d'analyse. Le graphique présente aussi le nombre total de moyens identifiés pour chaque catégorie. Parmi ceux-ci, pour chaque catégorie du modèle, le nombre de moyens spécifiques correspond au nombre de moyens qui réfèrent spécifiquement et uniquement à cette catégorie.

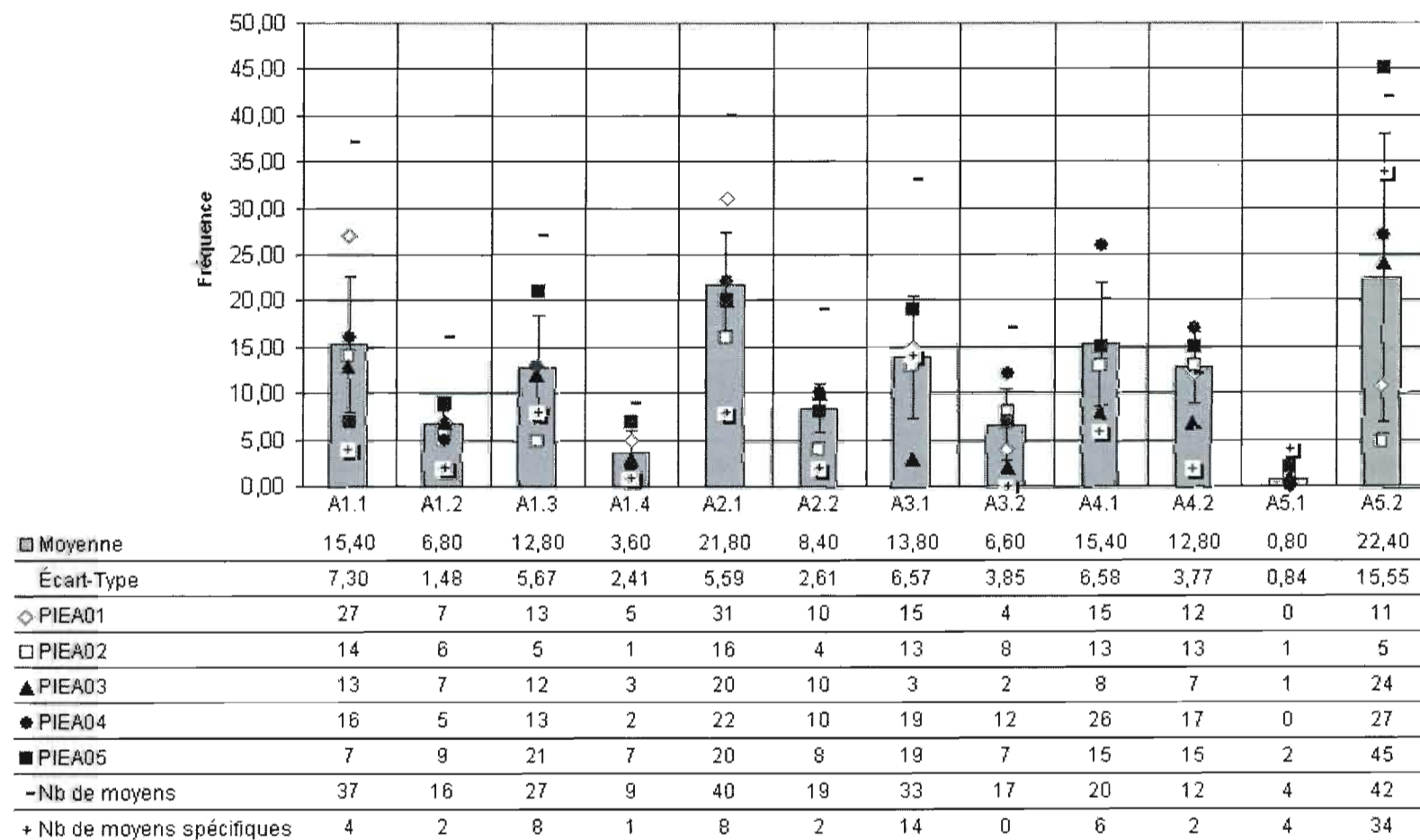


Figure 3.1 Répartition des énoncés en fréquences et nombre de moyens identifiés selon les catégories du modèle d'argumentation de la validité des inférences d'évaluation

La première ligne du tableau sous le graphique présente la moyenne du nombre d'énoncés catégorisés dans chaque argument de la structure d'argumentation de la validité des inférences d'évaluation, toutes politiques (PIEA) confondues. La deuxième ligne présente l'écart type associé à chacune de ces moyennes, c'est-à-dire la dispersion autour de la moyenne du nombre d'énoncé catégorisés dans chaque argument pour chaque politique (PIEA) analysée. Un écart type faible signifie que, pour chaque politique (PIEA) analysée, le nombre d'énoncés catégorisés dans l'argument identifié se situe près de la moyenne. On peut alors conclure que le nombre d'énoncés référant à l'argument correspondant est sensiblement homogène pour l'ensemble des politiques (PIEA) analysées. Dans le cas contraire, c'est-à-dire lorsque l'écart type est élevé, lorsque les données sont plus dispersées, on peut conclure que les politiques (PIEA) analysées diffèrent quant au nombre d'énoncés qui réfèrent à un argument donné. Les lignes trois à sept (PIEA01 à PIEA05) présentent, pour chaque politique (PIEA) analysée, le nombre d'énoncés catégorisés dans chaque argument de la structure d'argumentation de la validité des inférences d'évaluation. Rappelons qu'un même énoncé pouvait être classé dans plus d'une catégorie d'analyse. Les données des lignes un à sept seront abordées dans la section 3.2 en réponse au premier objectif de ce mémoire portant sur la validation la structure d'argumentation interprétative de Kane (2006).

La ligne huit, avant-dernière ligne du tableau, présente, pour chaque politique (PIEA) analysée, le nombre total de moyens d'application et/ou d'évaluation identifiés pour

chaque argument de la structure d'argumentation de la validité des inférences d'évaluation. Un même moyen ayant plus d'une fonction pouvait être identifié pour plus d'un argument. La ligne neuf, dernière ligne du tableau, présente, pour chaque politique (PIEA) analysée, le nombre de moyens d'application et/ou d'évaluation spécifiques à chaque argument. Il s'agit de moyens ne s'appliquant qu'à l'argument identifié et à aucun autre argument de la structure d'argumentation. Ces moyens sont identifiés à l'appendice B et présentés à la section 3.3 en réponse au second objectif de ce mémoire.

3.2 Validation de la structure d'argumentation interprétative de Kane (2006)

Une analyse quantitative de statistiques descriptives (fréquences absolues et relatives, moyennes, écarts types) est effectuée à partir des données issues de l'analyse de contenu pour évaluer l'exhaustivité, l'exclusivité et la pertinence des catégories du modèle original de Kane (2006) dans le contexte de l'évaluation des apprentissages en enseignement collégial.

3.2.1 Exhaustivité des catégories du modèle de Kane (2006)

Les catégories doivent être exhaustives en ce sens qu'elles doivent pouvoir permettre de classer tous les éléments du matériel analysé. (L'Écuyer, 1987) L'analyse de la

fréquence des énoncés classés et non classés dans les catégories du modèle original de Kane (2006) permet de faire une évaluation de leur exhaustivité au regard du matériel analysé. Comme il a été mentionné, le processus de catégorisation à l'étape de la modélisation correspond à un modèle mixte (L'Écuyer, 1987, p. 59) permettant d'effectuer des regroupements, des subdivisions ou des ajouts au modèle de Kane (2006) si celui-ci ne couvre pas l'ensemble du contenu analysé. Le tableau 3.1 présente les résultats de la réduction des données à travers le processus de catégorisation.

Tableau 3.1 Réduction des données à travers le processus de catégorisation

Processus de catégorisation	Fréquence (Nb d'énoncés)	Fréquence relative
1- Identification des énoncés significatifs	424	100 %
2- Catégorisation dans le modèle de Kane (2006) (A1.1 à A4.2)	268	63,21 %
<i>Inclassables</i>	<i>156</i>	<i>36,79 %</i>
3- Catégorisation dans les deux catégories émergentes (A5.1 et A5.2)	116	27,36 %
<i>Inclassables</i>	<i>40</i>	<i>9,43 %</i>

268 énoncés (63,21%) ont été classés dans les catégories du modèle de Kane (2006). 156 énoncés (36,79%) n'ont pas pu être classés dans les catégories du modèle de Kane (2006). De l'analyse de ces derniers sont émergées deux nouvelles catégories : deux arguments ajoutés au modèle initial. Ces deux arguments (A5.1 et A5.2) font référence à la crédibilité de l'inférence d'évaluation. Après l'analyse de ces 156 énoncés en fonction des deux catégories ajoutées, seulement 40 énoncés sont demeurés inclassables dans les catégories du modèle, faisant passer le pourcentage d'énoncés non considérés par le modèle de 36,79% à 9,43%. Ces 40 énoncés restants sont soit des énoncés généraux sur l'évaluation des apprentissages, qui n'apportent pas de contribution particulière à l'argumentation de la validité des inférences d'évaluation, soit des pratiques incohérentes avec l'argumentation de la validité de l'inférence d'évaluation. Il en sera discuté davantage au chapitre IV.

La nécessité d'ajouter des catégories au modèle de Kane (2006) indique que les arguments de validité du modèle initial ne couvrent pas l'ensemble du phénomène de l'évaluation des apprentissages dans les établissements d'enseignement collégial. Pour y remédier, deux catégories supplémentaires sont suggérées en complément au modèle de Kane (2006) pour assurer une représentation plus exhaustive de l'argumentation de la validité de l'évaluation des apprentissages dans les établissements d'enseignement collégial.

3.2.2 Exclusivité des catégories du modèle de Kane (2006)

L'exclusivité des catégories fait référence au fait qu'un même élément ne puisse être classé dans deux catégories différentes. La nécessité d'exclusivité est toutefois contestée puisque un même énoncé peut parfois renfermer plus d'un sens. (L'Écuyer, 1987) C'est d'ailleurs ce que révèlent les résultats de la présente recherche par rapport aux catégories du modèle de Kane (2006). Le tableau 3.2 présente la répartition des énoncés selon le nombre de catégories dans lesquels ils ont été classés.

Tableau 3.2 Répartition des énoncés selon le nombre de catégories dans lesquels ils ont été classés

	Nombre de catégories										Total
	0	1	2	3	4	5	6	7	8	9	
Fréq. (Nb d'énoncés)	40	239	74	27	13	2	25	1	1	2	424
Fréq.rel.	9,43	56,37	17,45	6,37	3,07	0,47	5,89	0,24	0,24	0,47	100

La majorité (56,37%) des énoncés analysés ont été classés dans une seule catégorie du modèle. 17,45% des énoncés ont été classés dans deux catégories à la fois. 16,75% des extraits ont été classés dans plus de deux catégories. 9,43% des extraits n'ont été classés dans aucune catégorie du modèle pour les raisons évoquées précédemment.

Les catégories du modèle ne sont pas totalement exclusives, mais on pourrait difficilement s'attendre à ce qu'elles le soient étant donné qu'une même pratique dans un établissement d'enseignement peut avoir plusieurs fonctions et ainsi répondre à plusieurs critères de validité.

3.2.3 Pertinence des catégories du modèle de Kane (2006)

Les catégories du modèle semblent pertinentes puisqu'un certain nombre d'énoncés se rapportent à chacune d'elles (voir la figure 3.1, p. 45). On pourrait toutefois remettre en question la pertinence de deux catégories, soit l'argument A5.1 dans lequel seulement quatre énoncés ont été classés et l'argument A1.4 dans lequel seulement 18 énoncés ont été classés. Ainsi, si un moins grand nombre d'énoncés ont été catégorisés comme faisant référence à ces arguments, il est possible que ces arguments soient moins pertinents dans le contexte de l'évaluation des apprentissages dans les établissements d'enseignement collégial ou dans les politiques institutionnelles d'évaluation des apprentissages (PIEA). L'argument du nombre n'est

toutefois pas suffisant pour remettre en question la pertinence des arguments du modèle. Il est possible que certains arguments du modèle soient pertinents pour assurer l'argumentation de la validité des inférences d'évaluation, mais que les pratiques liées à ceux-ci soient moins connues ou moins répandues que d'autres.

3.2.4 Synthèse des résultats du premier objectif

L'analyse des politiques institutionnelles d'évaluation des apprentissages (PIEA) du réseau collégial québécois en fonction de la structure d'argumentation interprétative de Kane (2006) révèle que l'ensemble de la structure s'applique à l'évaluation des apprentissages à l'enseignement collégial. Toutefois, la structure d'argumentation de Kane (2006) ne couvre pas l'ensemble de la réalité de l'évaluation des apprentissages à l'enseignement collégial. C'est pourquoi un niveau comprenant deux arguments supplémentaires est suggéré en complément au modèle initial. Il s'agit du niveau *Crédibilité* qui réfère à la nécessité pour que l'inférence puisse être considérée valide qu'elle soit acceptée par les acteurs impliqués. Les acteurs impliqués comprennent autant les personnes de l'établissement impliquées dans la production ou l'interprétation de l'inférence d'évaluation (étudiants, professeurs et autres membres de l'établissement) que des personnes extérieures à l'établissement qui elles aussi interprètent les résultats de l'évaluation (employeurs, membres d'autres établissements d'enseignement, population générale). Le fait que l'ensemble des

arguments soit respecté de façon à ce que chacune des inférences considérées de façon individuelle soit convaincante (A5.1) et le fait que les processus d'apprentissage et d'évaluation soient connus des acteurs (A5.2) facilitent la crédibilité de l'inférence d'évaluation auprès des acteurs impliqués. La figure 3.2 présente le modèle issu des résultats de la présente recherche : une adaptation de la structure d'argumentation de Kane (2006), suggérée pour l'interprétation des inférences d'évaluation dans le contexte de l'évaluation des apprentissages au collégial.

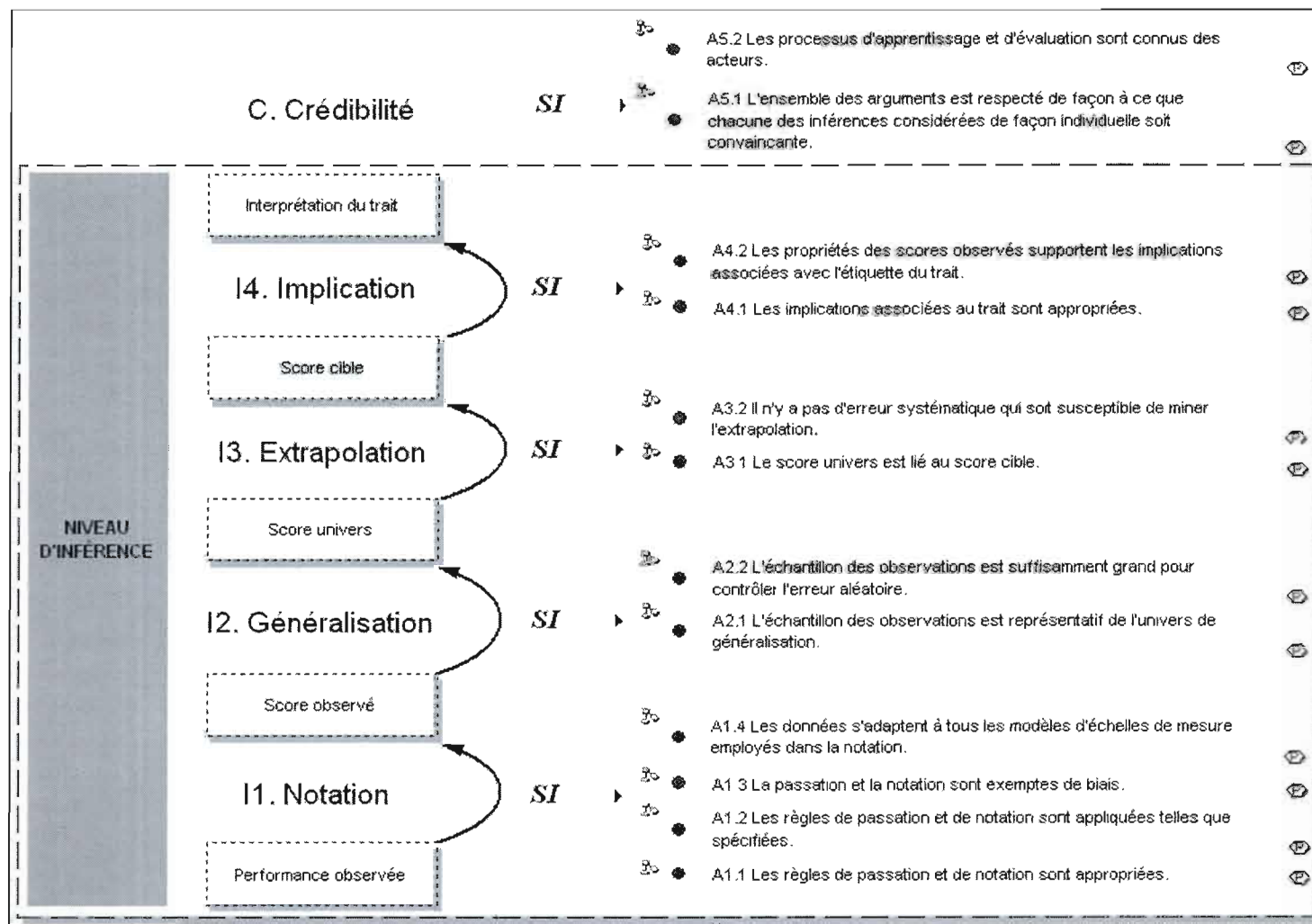


Figure 3.2 Structure d'argumentation de la validité des inférences d'évaluation des apprentissages

3.3 Identification de moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages

Le second objectif de la présente recherche est d'identifier des moyens que peuvent mettre en place les institutions pour chercher à assurer la validité des inférences d'évaluation au regard des apprentissages des étudiants. Les analyses réalisées ont permis d'identifier un certain nombre de moyens présentés de façon détaillée à l'appendice B. Les moyens identifiés sont issus des politiques institutionnelles d'évaluation des apprentissages (PIEA) analysées. Une synthèse des résultats les plus significatifs pour chaque argument du modèle d'argumentation de la validité des inférences d'évaluation des apprentissages est présentée dans cette section.

3.3.1 Inférence de notation (I1)

L'inférence de notation (figure 3.3) est réalisée à travers la passation et la notation des activités d'évaluation de façon à pouvoir faire le passage de l'interprétation de la performance observée au score observé. La performance est alors interprétée en termes de score associé à un échantillon d'observations constituant la mesure.

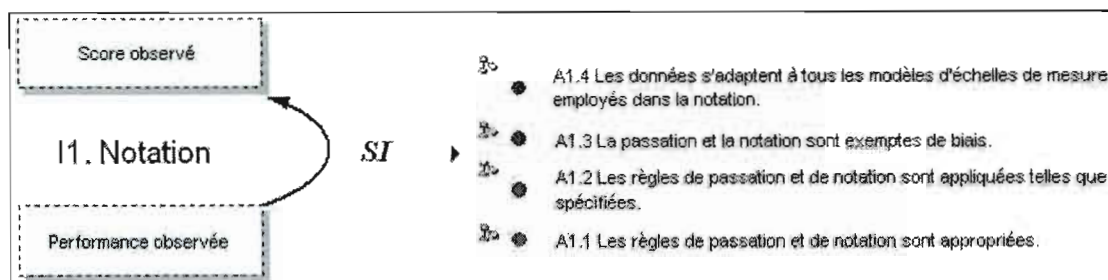


Figure 3.3 Inférence de notation (I1)

3.3.1.1 Les règles de passation et de notation sont appropriées (A1.1)

Les moyens identifiés dans les politiques institutionnelles d'évaluation des apprentissages (PIEA) relativement à l'argument A1.1 concernent la planification des activités d'évaluation pour chercher à s'assurer que la passation et la notation soient appropriées.

Un premier moyen pour chercher à appliquer cet argument est la tenue d'activités de concertation entre les professeurs d'un département. Ensembles, les professeurs d'un département constituent un groupe d'experts du domaine cible. Leur concertation au sujet des objectifs, des contenus, des modes d'évaluation, des exigences et des seuils de réussite des cours et des programmes qui les concernent est un moyen de chercher à s'assurer que les règles de passation et de notation sont appropriées. La concertation des professeurs permet aussi de rechercher l'équivalence des objectifs d'apprentissage, des procédés et des critères d'évaluation dans un même cours

dispensé par plusieurs enseignants. C'est donc un moyen de chercher à s'assurer que la correction des activités d'évaluation sommative est réalisée à partir de critères établis par l'ensemble des enseignants qui donnent le même cours et selon une pondération équivalente.

En établissant et en diffusant des règles départementales d'évaluation des apprentissages, le département peut chercher à s'assurer que les règles de passation et de notation sont appropriées, mais aussi qu'elles tiennent compte des particularités de chaque programme d'étude tout en assurant une certaine uniformité des modalités d'évaluation au sein d'un même département contribuant à la recherche d'équivalence. Ces règles sont approuvées au sein du département et par la direction des études. L'élaboration de plans de cours cadre auxquels les professeurs peuvent se référer lors de la planification de leurs cours et de leurs activités d'évaluation est un moyen pour le département de répondre à ces mêmes objectifs. Le professeur doit quant à lui rédiger un plan de cours conforme à la politique institutionnelle d'évaluation des apprentissages (PIEA), aux règles départementales d'évaluation des apprentissages, au programme et au plan de cours cadre dans lequel il établit les modalités d'évaluation : la nature des épreuves (travaux, examens ou autres) servant à l'établissement de la note finale, leur pondération, leur durée et la date prévue. Le plan de cours et ses modifications en cours de session sont approuvés par le département et les modalités d'analyse et d'approbation des plans de cours sont elles-mêmes approuvées par la direction des études. Toutes étapes d'approbation

constituent des consultations de différents groupes experts qui contribuent à la validité de l'argument interprétatif.

Enfin, l'organisation d'activités de perfectionnement tout en offrant des services d'encadrement aux professeurs permet de tenir ceux-ci informés des derniers développements du domaine de la mesure et de l'évaluation et de leur donner les outils et les méthodes nécessaires pour chercher à s'assurer que les règles de passation et de notation sont appropriées.

3.3.1.2 Les règles de passation et de notation sont appliquées telles que spécifiées (A1.2)

L'établissement de règles de passation et de notation appropriées contribue à assurer la validité de l'inférence de notation, mais encore faut-il que celles-ci soient appliquées telles que spécifiées.

Concernant cet argument, la concertation entre les professeurs du département au sujet des objectifs, des contenus, des modes d'évaluation, des exigences et des seuils de réussite des cours et des programmes qui les concernent est aussi importante. Si les règles de passation et de notation sont discutées lors des activités de concertation

et que les professeurs sont impliqués, s'ils ont participé à leur élaboration, on peut supposer qu'ils auront davantage tendance à les appliquer telles que spécifiées.

Les professeurs devraient aussi s'assurer de connaître et d'appliquer la politique institutionnelle d'évaluation des apprentissages (PIEA) et les règles départementales d'évaluation des apprentissages lors de la passation et la notation des activités d'évaluation dont ils assument eux-mêmes la responsabilité. L'évaluation est alors fondée sur l'atteinte des objectifs d'apprentissage annoncés dans le plan de cours ou sur l'atteinte des éléments de la compétence, selon les standards de réussite prescrits ou définis par le département.

Les services d'encadrement et de perfectionnement contribuent aussi à s'assurer que les professeurs connaissent et appliquent la politique institutionnelle d'évaluation des apprentissages (PIEA) et les règles départementales d'évaluation des apprentissages. Cela en supposant que les activités d'encadrement et de perfectionnement améliorent les pratiques.

Enfin, la direction des études peut, en plus de décrire et de diffuser les procédures administratives requises pour la passation et la notation des activités d'évaluation, évaluer les pratiques d'évaluation au sein du collège et rédiger un rapport écrit sur le processus d'évaluation des apprentissages qui soit remis au conseil d'administration

du collège à la fin de chaque session ou de chaque année. Ce rapport ferait entre autres état du degré d'application des règles de passation et de notation.

3.3.1.3 La passation et la notation sont exemptes de biais (A1.3)

L'inférence de notation peut être affectée par certaines sources de biais liées aux caractéristiques de l'évaluateur ou de l'étudiant ou à des caractéristiques environnementales. Le contenu des politiques institutionnelles d'évaluation des apprentissages (PIEA) contient un certain nombre de moyens permettant de réduire ces sources de biais lors de la passation ou de la notation.

Les étudiants qui suivent un même cours devraient être évalués à partir des mêmes objectifs, des mêmes critères et selon une même pondération, quel que soit le professeur qui donne le cours. La concertation entre les professeurs d'un même département et entre les départements au sujet des modes et des instruments d'évaluation et à propos du degré d'exigence et des seuils de réussite est un moyen de rechercher l'équivalence des évaluations pour réduire les risques de biais liés à l'évaluateur. Il est aussi suggéré d'avoir un seul et même plan de cours pour un cours donné à plusieurs groupes d'étudiants, à la même session dans un même secteur, ce qui permet de s'assurer que le contenu enseigné et les activités d'évaluation sont les mêmes pour tous, et donc assure une évaluation plus équitable. Dans une des

politiques institutionnelles d'évaluation des apprentissages (PIEA) analysées, il est même suggéré d'avoir un seul et même examen final pour un cours donné à plusieurs groupes d'étudiants, à la même session dans un même secteur. En plus de réduire les risques de biais liés à l'évaluateur, cela facilite la reconnaissance des acquis lors de l'évaluation des équivalences. Il sera davantage question de l'évaluation des équivalences et des substitutions au chapitre IV.

Un autre moyen de réduire les risques de biais lors de la notation est l'établissement par le professeur de critères et de standards clairs préalablement à la passation et à la notation. En se référant lors de la notation aux critères et aux standards préétablis, le professeur réduit les risques de glissement lors de la notation et assure ainsi une évaluation plus équitable. Certains biais de notation peuvent aussi être liés au manque de pertinence des critères d'évaluation.

Le fait de faire porter l'inférence d'évaluation sur un seul score ou une seule activité d'évaluation pourrait être une source de biais liée aux conditions d'évaluation, à l'environnement ou à l'état physique ou psychologique de la personne évaluée au moment de la passation. Un moyen de réduire ces sources de biais peut être de répartir l'évaluation sur plusieurs activités en cours de session. La direction des études peut aussi structurer le calendrier des évaluations pour l'ensemble du collège de façon à chercher à s'assurer qu'aucun étudiant n'ait plus de deux évaluations par

jour et s'assurer que les conditions environnementales d'évaluation sont adéquates (ex. : grandeur et ameublement des locaux, lumière, contrôle des bruits environnants).

Une attention particulière devrait être accordée par le professeur, notamment lors de l'évaluation de travaux d'équipe, à chercher à s'assurer que la notation porte sur l'évaluation des apprentissages individuels et qu'elle ne soit pas biaisée par la performance des autres étudiants.

L'étudiant a lui-même une responsabilité au regard de la réduction des risques de biais de notation en s'assurant de connaître et d'appliquer la politique institutionnelle d'évaluation des apprentissages (PIEA) et les règles départementales d'évaluation des apprentissages. Par exemple, un étudiant qui contrevient aux règles d'évaluation des apprentissages en trichant ou en plagiant lors d'une évaluation introduit un biais réduisant la validité de l'inférence de notation puisque celle-ci ne se rapporte pas à sa propre performance. À ce propos, le professeur peut réduire les risques de triche en s'assurant de la confidentialité des instruments d'évaluation avant leur administration aux étudiants. Un autre exemple de responsabilité de l'étudiant est celle de respecter les normes de présentation lors de la rédaction de ses travaux écrits, ce qui peut réduire les risques que la notation de ses réalisations soit biaisée par leur présentation matérielle. Quant à l'enseignant, il a la responsabilité de fournir aux étudiants les consignes de présentation des travaux.

3.3.1.4 Les données s'adaptent à tous les modèles d'échelles de mesure employés dans la notation (A1.4)

Les politiques institutionnelles d'évaluation des apprentissages (PIEA) n'abordent pratiquement pas la question des échelles de mesure. Une seule politique institutionnelle d'évaluation des apprentissages (PIEA) y fait directement référence dans deux énoncés. C'est à partir de ceux-ci qu'a pu être formulé un moyen qui se rapporte spécifiquement à cet argument : celui d'établir clairement l'échelle de notation employée et la façon dont les cotes sont interprétées numériquement.

Les autres énoncés ne sont qu'indirectement reliés à l'argument, c'est pourquoi les moyens tirés des politiques institutionnelles d'évaluation des apprentissages (PIEA) concernant cet argument sont d'ordre plus général. Ils concernent le développement des instruments de mesure et l'établissement des seuils de réussite sans aborder spécifiquement les préoccupations liées aux données de l'évaluation et aux échelles de mesures.

3.3.2 Inférence de généralisation (I2)

L'inférence de généralisation (figure 3.4) est réalisée à travers l'échantillon des observations. C'est la constitution de l'échantillon des observations qui assure la validité de l'inférence de généralisation qui permet le passage de l'interprétation du score observé au score univers. Le score est alors interprété non plus seulement en lien avec la mesure, mais à travers sa représentativité d'un sous-ensemble plus large d'observations que constituent le cours et le programme d'étude.

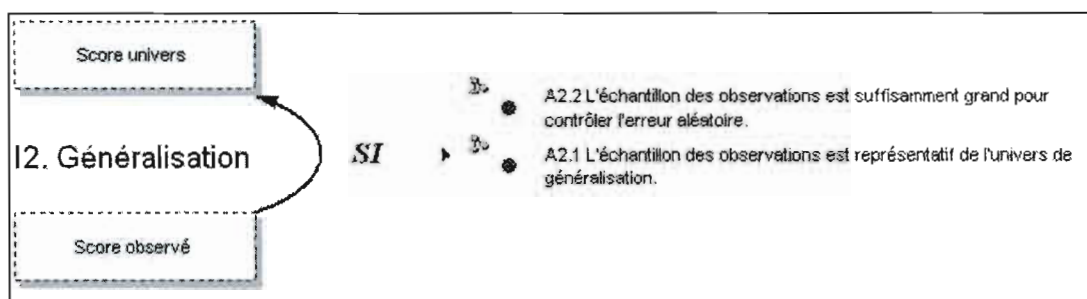


Figure 3.4 Inférence de généralisation (I2)

3.3.2.1 L'échantillon des observations est représentatif de l'univers de généralisation (A2.1)

C'est à l'argument A2.1 du modèle original de Kane (2006) que le plus grand nombre d'énoncés analysés réfèrent et dans lequel le plus grand nombre de moyens d'application ou d'évaluation ont été identifiés, sans toutefois avoir le plus grand nombre de moyens spécifiques à l'argument. Cela signifie qu'il s'agit d'une catégorie d'ordre plus général. Cet argument réfère à la représentativité de l'échantillon des observations par rapport à l'univers de généralisation, donc à la représentativité des items ou des observations qui forment le cours et la mesure par rapport au programme d'étude.

Le fait de planifier les activités d'apprentissage et d'évaluation en conformité avec les plans cadres ou avec les objectifs et standards de chacun des cours est un moyen pour le professeur de chercher à s'assurer que l'échantillon des observations est représentatif de l'univers de généralisation. La planification de l'évaluation permet de s'assurer que les contenus, les objectifs et les critères d'évaluation sont définis, clairement présentés et établis dès le départ. Cette planification peut être faite conjointement à la rédaction du plan de cours qui présente : la compétence visée par le cours et l'apport du cours à la réalisation des objectifs du programme ; les objectifs de formation du cours présentés en énoncés de compétence et en standards et les objectifs du cours ; les contenus couverts par le cours ; les activités d'enseignement et

d'apprentissage ; les activités d'évaluation sommative des apprentissages et leurs modalités, c'est-à-dire la nature des épreuves (travaux, examens ou autres) servant à l'établissement de la note finale, leur pondération, leur durée et la date prévue des évaluations. La conception des activités d'évaluation et le développement des instruments de mesure sont fondés sur les objectifs du cours ou les compétences à développer. Raïche (2008, p. 7) suggère d'ailleurs, en plus de la rédaction d'un plan de cours, la rédaction d'un plan d'évaluation qui présente les objectifs et les modalités d'évaluation plutôt que les activités d'enseignement et les objectifs d'apprentissage.

Le fait de répartir l'évaluation sur plusieurs activités d'évaluation en cours de session permet aussi d'assurer une meilleure représentativité de l'univers de généralisation en augmentant le nombre d'observations et en diminuant les risques de biais liés aux conditions d'évaluation, à l'environnement ou à l'état physique ou psychologique de la personne évaluée au moment de l'évaluation.

Le département peut s'assurer de la concertation des enseignants sur les objectifs, les contenus, les modes d'évaluation, les exigences et les seuils de réussite des cours et des programmes qui les concernent. Ainsi, les professeurs peuvent se concerter pour s'assurer que les items ou les observations qui forment la mesure sont représentatifs du cours et du programme. Non seulement cela permet de fonder le contenu de l'évaluation sur le jugement d'une série d'experts, mais cela permet aussi la recherche

d'équivalence des objectifs d'apprentissage, des procédés et des critères d'évaluation dans un même cours dispensé par plusieurs enseignants et même dans ce cas, permettre l'élaboration d'instruments d'évaluation communs.

Le département peut élaborer des règles d'évaluation propres au programme qui précisent comment la maîtrise de la compétence est évaluée pour obtenir la note de passage et élaborer un plan cadre pour chacun des cours dont il est responsable. Le plan cadre présente les contenus essentiels du cours, ce qui peut guider les professeurs dans la détermination des contenus devant faire l'objet de l'évaluation. De plus, en se renseignant sur les besoins de formation des enseignants et en organisant, en collaboration avec le service de formation continue et la direction des études, des services d'encadrement et de perfectionnement, le département peut chercher à s'assurer que les enseignants disposent des méthodes et des outils nécessaires au développement d'instruments de mesure qui permettent de générer des résultats représentatifs de la compétence de l'étudiant dans le cours et le programme d'études.

Le comité de programme quant à lui est responsable de la détermination du contenu des cours en lien avec les objectifs et les standards du programme, du domaine et des exigences du collège et du Ministère. Il peut procéder, en collaboration avec le département et les enseignants, à l'évaluation des instruments et des méthodes d'évaluation afin de s'assurer qu'ils sont conformes à la politique institutionnelle d'évaluation des apprentissages (PIEA), aux règles départementales d'évaluation des apprentissages, au programme et aux plans de cours cadre.

3.3.2.2 L'échantillon des observations est suffisamment grand pour contrôler l'erreur aléatoire (A2.2)

Dans la planification de l'évaluation, un moyen pour le professeur de chercher à s'assurer que l'échantillon des observations est suffisamment grand pour contrôler l'erreur aléatoire est d'attribuer une pondération suffisamment importante aux activités d'évaluation pour qu'elle soit significative de la maîtrise globale des éléments de compétences, de la ou des compétences visées par le cours. Le département peut aussi prescrire dans le plan cadre une pondération suffisamment importante des activités d'évaluation pour être significative de la maîtrise globale des éléments de compétences, de la ou des compétences visées par le cours. Le fait de répartir l'évaluation sur plusieurs activités d'évaluation en cours de session permet aussi de contrôler l'erreur aléatoire en augmentant le nombre d'observations et en

diminuant les risques de biais liés aux conditions d'évaluation, à l'environnement ou à l'état physique ou psychologique de la personne évaluée au moment de l'évaluation.

3.3.3 Inférence d'extrapolation (I3)

C'est au niveau de l'inférence d'extrapolation (figure 3.5) qu'entrent en considération les questions relatives à la définition du programme d'étude. Le score est alors interprété en lien avec le programme d'étude et le domaine cible. Les liens doivent alors être clairement établis entre le programme d'étude et le domaine, car c'est par l'extrapolation qu'on peut émettre le jugement affirmant que l'étudiant évalué possède ou ne possède pas les compétences du domaine cible.

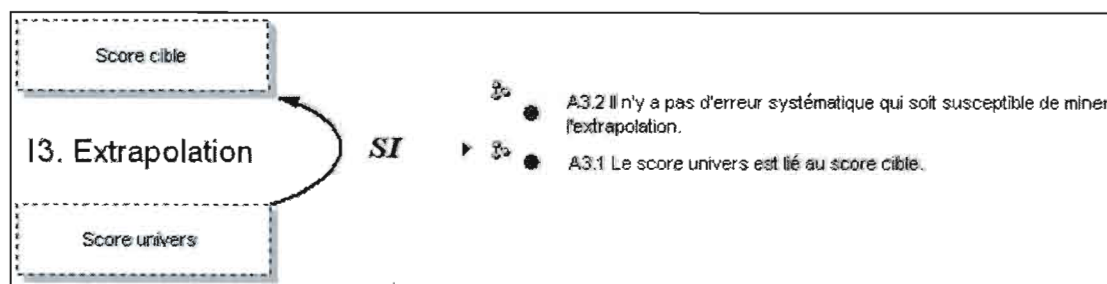


Figure 3.5 Inférence d'extrapolation (I3)

3.3.3.1 Le score univers est lié au score cible (A3.1)

Pour l'argument A3.1, le résultat est interprété quant à sa représentativité par rapport au domaine cible. Les moyens identifiés pour appliquer ou évaluer cet argument de validité sont donc des moyens permettant d'établir un lien entre le programme et le domaine.

Le professeur planifie les activités d'apprentissage et d'évaluation en conformité avec la politique institutionnelle d'évaluation des apprentissages (PIEA), les règles départementales d'évaluation des apprentissages, le programme et le plan de cours cadre. Un service de soutien pédagogique assure un rôle d'assistance aux professeurs, et particulièrement aux nouveaux professeurs, dans la préparation et l'organisation des cours en accord avec les objectifs, les standards et les politiques du programme et du département. Le professeur indique dans le plan de cours la place du cours dans le programme d'études. Ce faisant, il établit le lien entre le cours, le programme et le domaine. En fournissant une bibliographie et des lectures complémentaires dans le plan de cours, le professeur établit le lien entre le cours et le domaine cible et incite les étudiants à faire de même. Le professeur démontre ainsi que son cours est constitué d'un sous-ensemble d'observations tiré du domaine cible et que d'autres éléments du domaine peuvent être explorés par l'étudiant lui-même, bien que ceux-ci n'aient pas été inclus dans l'univers de généralisation (ensemble des observations abordées dans le programme et plus spécifiquement dans cours et susceptibles de

faire partie de l'évaluation). Lors de la planification de l'évaluation, le professeur centre l'évaluation sur les objectifs du programme.

Le département et le comité de programme sont responsables de l'élaboration des objectifs du programme. Le comité de programme élabore les critères spécifiques aux compétences professionnelles du domaine et développe les critères et les standards d'évaluation pour le programme. Il assure la concordance entre les éléments de la politique institutionnelle d'évaluation des apprentissages (PIEA), le développement de programme et les plans de cours. Il cherche aussi à favoriser l'harmonisation des pratiques et l'équivalence des évaluations pour tous les cours du programme. L'harmonisation des pratiques favorise le lien entre le score univers et le score cible. Chaque cours du programme constitue en quelque sorte un univers de généralisation duquel sont tirés les échantillons d'observations qui constituent chaque évaluation. L'harmonisation des pratiques et l'équivalence des évaluations favorisent l'établissement de liens entre les cours, donc entre différents univers de généralisation et offre une meilleure représentativité du domaine cible. Le département s'assure que les cours sont intégrés au programme et qu'ils contribuent à l'atteinte des objectifs du programme. Les liens entre un cours et les autres cours du programme sont présentés dans le plan cadre de façon à bien établir les liens entre les différents univers de généralisation assurant une meilleure représentativité du domaine cible.

L'évaluation des équivalences et des substitutions équivaut à vérifier l'équivalence de deux univers de généralisation au sein d'un même domaine cible ou même parfois de domaines différents. La direction des études doit dans ces cas évaluer si le score univers est lié au score cible sans toutefois être en mesure d'évaluer la validité des arguments liés à la notation et à la généralisation, qui sont liées aux pratiques d'évaluation réalisées dans le cours qui sert d'équivalence ou de substitution et qui ne sont donc plus disponibles pour permettre une vérification de la validité des inférences réalisées. Un moyen suggéré dans les politiques institutionnelles d'évaluation des apprentissages (PIEA) pour évaluer les équivalences et les substitutions est d'élaborer et d'administrer un examen démontrant la maîtrise par l'étudiant de la compétence et des éléments de compétence prévus au cours faisant l'objet de la demande.

C'est ici, au niveau de l'extrapolation, que l'épreuve synthèse de programme prend toute son importance. Elle permet d'attester de l'intégration des apprentissages réalisés dans l'ensemble du programme. L'épreuve synthèse permet donc d'établir le lien entre les apprentissages de l'étudiant et le domaine cible. Le comité de programme élabore l'épreuve synthèse sur la base des compétences, des objectifs et des standards du programme d'études et non sur le contenu spécifique de chacun des cours du programme. L'épreuve synthèse tient compte des objectifs et des standards déterminés dans le programme, du profil de sortie de l'étudiant et de la recherche d'harmonisation et d'équivalence.

Avant de recommander au Ministère de décerner le diplôme à un étudiant, et ainsi attester de la compétence de cet étudiant dans le domaine cible, les instances du collège vérifient si toutes les conditions de validation de l'argument interprétatif établies par le Ministère et par le collège ont été rencontrées.

3.3.3.2 Il n'y a pas d'erreur systématique qui soit susceptible de miner l'extrapolation (A3.2)

Pour appliquer l'argument A3.2, il faut chercher à identifier et à contrôler les plus sérieuses sources d'erreurs systématiques qui pourraient biaiser l'inférence d'extrapolation, c'est-à-dire l'interprétation du résultat relativement à la compétence dans le domaine cible.

Lors de la conception des modalités d'évaluation, lors du choix de la méthode d'évaluation (ex. : questions objectives, questions à développement, performances), des règles de notations et des conditions d'observation, la concertation des professeurs au sujet des modes et des instruments d'évaluation permet d'évaluer, d'identifier et de contrôler les plus sérieuses sources d'erreurs systématiques qui pourraient être susceptibles de miner l'extrapolation de l'interprétation de l'univers de généralisation au domaine cible.

Les politiques institutionnelles d'évaluation des apprentissages (PIEA) font très peu référence aux propriétés de la mesure et à la construction des instruments de mesure. La source d'erreur la plus souvent mentionnée est le niveau de langue parlée et écrite. En effet, l'évaluation des apprentissages d'un étudiant peut être biaisée s'il ne possède pas les compétences linguistiques requises par la méthode d'évaluation. L'étudiant a sa part de responsabilité à cet égard et il devrait chercher à s'assurer que sa performance ne soit pas biaisée par certaines compétences auxiliaires, telles que ses compétences linguistiques. Il peut donc se renseigner sur les exigences linguistiques selon lesquelles il sera évalué et développer, durant ses études collégiales, les compétences linguistiques requises à son niveau d'études. Le professeur, quant à lui, peut détecter les sources d'erreur systématique et référer l'étudiant éprouvant des difficultés vers une ressource d'aide appropriée. S'il s'agit des compétences linguistiques, il peut par exemple référer l'étudiant au centre d'aide en français (ou en anglais selon le cas) du collège.

3.3.4 Inférence d'implication (I4)

L'inférence d'implication (figure 3.6) consiste à interpréter le score en lien avec les implications associées avec la compétence et le domaine cible. Il s'agit de considérer dans l'interprétation du score les contextes dans lesquels la ou les compétences évaluées sont impliquées, les contextes académiques et professionnels, les contextes de mesure et du quotidien. C'est dans l'inférence d'implication que sont aussi considérées les exceptions et conditions particulières pouvant influencer l'interprétation.

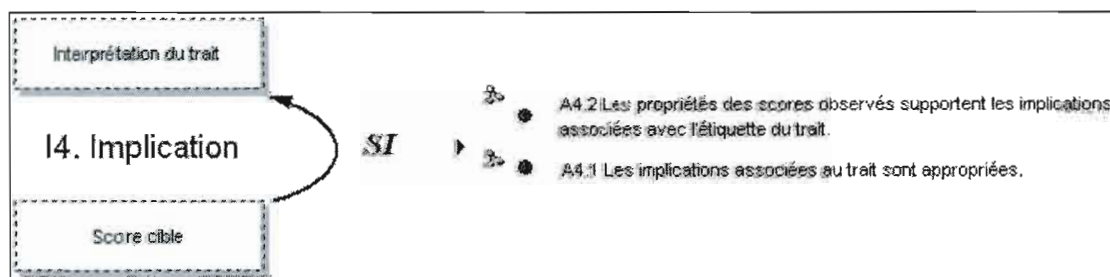


Figure 3.6 Inférence d'implication (I4)

3.3.4.1 Les implications associées au trait sont appropriées (A4.1)

Les implications regroupent les extrapolations de l'interprétation pour y inclure toutes affirmations ou suggestions associées au trait ou à la compétence cible. Les implications concernent donc les inférences réalisées du résultat au domaine cible en passant par les liens établis entre le programme d'étude et le domaine. C'est pourquoi le comité de programme doit définir le profil de sortie de l'étudiant inscrit au programme en lien avec le domaine cible. Le conseil d'administration, quant à lui, établit la liste des conditions que doit respecter l'étudiant pour être diplômé de son programme, comme par exemple avoir rencontré tous les objectifs du programme et avoir réussi l'épreuve synthèse et les épreuves ministérielles. L'établissement, en recommandant au Ministère de décerner un diplôme à un étudiant, affirme que ce dernier, en ayant répondu à chacune de ces conditions, devrait être compétent dans le domaine. Les conditions mentionnées concernant la sanction des études et la remise des diplômes constituent des implications associées aux traits ou aux compétences du domaine pour lequel l'étudiant est diplômé. Ainsi, les interprétations associées aux compétences maîtrisées sont bien justifiées par rapport au domaine cible.

Les implications concernent aussi toutes les exceptions appliquées au score et à son interprétation, c'est-à-dire toutes conditions particulières pouvant affecter le résultat qui ne sont pas nécessairement reliées à la compétence dans le domaine cible. Il s'agit par exemple des exigences liées à la présence aux cours ou aux activités d'évaluation,

à la présentation matérielle des travaux ou à la qualité de la langue. Tous ces exemples font référence à des compétences auxiliaires dont l'établissement considère qu'elles doivent nécessairement être maîtrisées pour que la compétence dans le domaine cible soit entièrement atteinte. La direction des études détermine ces exceptions, le département les spécifie dans leurs règlements et le professeur en informe les étudiants et les précise dans le plan de cours. La pertinence de l'évaluation de certaines compétences auxiliaires est abordée au chapitre IV.

3.3.4.2 Les propriétés des scores observés supportent les implications associées avec l'étiquette du trait (A4.2)

L'explication des choix liés à la pondération contribue à s'assurer que les propriétés des scores observés supportent les implications associées le trait ou la compétence cible. L'interprétation du score sera différente selon qu'on choisit une pondération plus importante de l'évaluation finale ou une répartition équivalente de la pondération des différentes activités d'évaluation. Les positions quant à ce choix varient entre les politiques institutionnelles d'évaluation des apprentissages (PIEA). Selon une première position, chaque cours comprend une évaluation finale dont la pondération est prédominante et plus élevée que chacune des évaluations antérieures. Si la pondération de l'évaluation finale n'est pas plus élevée que chacune des évaluations antérieures, l'étudiant doit obtenir la note de passage à cette évaluation finale pour

réussir le cours. L'évaluation finale est l'occasion pour l'étudiant de démontrer qu'au terme du cours, il maîtrise la compétence visée par le cours. Alors que l'étudiant est en cours d'apprentissage lors des évaluations en cours de session, il devrait en fin de session maîtriser la compétence visée et le démontrer au moyen de sa performance à l'évaluation finale. Le score permet alors d'attester de la maîtrise des compétences visées par le cours et supporte donc les implications liées au trait ou à la compétence cible. Selon une autre position, aucune activité d'évaluation ne devrait excéder, par exemple, 25% de la note globale d'un cours. Cette position est elle aussi justifiable par rapport à l'argument A4.2, puisqu'en situation authentique, la compétence se démontre rarement en une occasion unique. En accordant une pondération équivalente à plusieurs évaluations différentes à plusieurs moments, dans divers contextes, l'évaluation de la compétence visée est plus authentique et plus représentative de la compétence en contexte réel. Ici encore, le score supporte les implications liées au trait ou à la compétence cible. Les deux positions étant justifiables par rapport à l'argumentation de la validité de l'inférence d'évaluation, le choix demeure à la discrétion des départements, en autant qu'il soit explicité.

Le fait que les scores et les mentions pouvant figurer au bulletin d'études collégiales, de même que les exceptions à la notation, soient explicités dans la politique institutionnelle d'évaluation des apprentissages (PIEA) contribue aussi à s'assurer que les propriétés des scores supportent les implications associées au trait ou à la compétence cible. On précise ainsi quelles interprétations peuvent être faites des

scores. Par exemple, le score associé à une réussite étant « RE », les propriétés du score supportent les implications liées au fait que l'étudiant satisfait aux exigences de l'épreuve uniforme de français ou de l'épreuve synthèse de programme. Le score associé à l'équivalence étant « EQ », les propriétés du score démontrent qu'il s'agit d'une équivalence et donc supportent les implications liées au fait qu'il s'agit d'une équivalence. Le score « EQ » en lui-même est une implication qui limite l'inférence en mentionnant que la validité de l'argument interprétatif n'a pu être vérifiée qu'en partie (inférence d'extrapolation seulement). Il sera davantage question de l'évaluation des équivalences et des substitutions au chapitre IV. La description des scores doit donc être complète et non équivoque. Les propriétés de quelques scores mentionnés dans les politiques institutionnelles d'évaluation des apprentissages (PIEA) analysées ont soulevé certains questionnements quant à la validité de l'inférence d'évaluation qui pourrait être réalisée. Il en sera question au chapitre IV.

3.3.5 Crédibilité de l'inférence d'évaluation (C)

Le niveau *Crédibilité* (figure 3.7) a été ajouté à la structure d'argumentation de Kane (2006) de façon à représenter de façon plus exhaustive la réalité de l'évaluation des apprentissages en enseignement collégial. Ce niveau fait référence au fait que l'inférence d'évaluation ne peut être jugée valide que si elle est acceptée par les acteurs impliqués. Selon Kane (2006), pour que l'argument interprétatif soit convaincant, chacune des inférences considérées de façon individuelle doit être convaincante. Un argument a donc été ajouté au modèle initial de Kane (2006) : *A5.1 L'ensemble des arguments est respecté de façon à ce que chacune des inférences considérées de façon individuelle soit convaincante*. De plus, pour que les acteurs acceptent comme valide l'inférence d'évaluation, ils doivent connaître les processus impliqués dans l'évaluation des apprentissages. Un second argument a été ajouté en ce sens : *A5.2 Les processus d'apprentissage et d'évaluation sont connus des acteurs*.

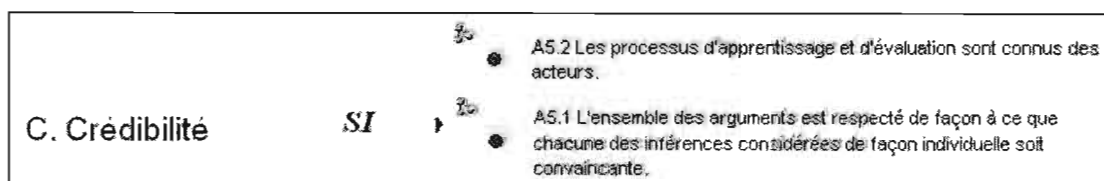


Figure 3.7 Crédibilité de l'inférence d'évaluation

3.3.5.1 L'ensemble des arguments est respecté de façon à ce que chacune des inférences considérées de façon individuelle soit convaincante (A5.1)

Pour que l'argument interprétatif soit convaincant, chaque inférence considérée de façon individuelle doit être convaincante. Une faille dans n'importe laquelle des inférences produites invalide l'argument interprétatif en entier, même si tous les arguments de validité des autres inférences sont respectés et que leur validité est démontrée (Crooks et *al.*, 1996 : voir Kane, 2006, p. 34).

Les moyens identifiés relativement à cet argument sont de nature méta-évaluative. Ils concernent l'évaluation des pratiques d'évaluation des apprentissages en accord avec la politique institutionnelle d'évaluation des apprentissages (PIEA) et les arguments de validité des inférences d'évaluation. Ainsi, la commission des études vérifie que les modalités départementales d'évaluation des apprentissages sont conformes à la politique institutionnelle d'évaluation des apprentissages (PIEA) et le département vérifie que les pratiques d'évaluation y sont conformes. Le comité de programme prend les mesures nécessaires lorsqu'il est avisé du non respect de la politique institutionnelle d'évaluation des apprentissages (PIEA). Et la direction des études établit un portrait statistique de l'évaluation des apprentissages pour chaque département ou programme et de l'application de chaque argument de validité à la fin de chaque année.

3.3.5.2 Les processus d'apprentissage et d'évaluation sont connus des acteurs (A5.2)

On peut penser que l'inférence d'évaluation sera plus facilement acceptée comme valide si les processus d'apprentissage et d'évaluation sont connus des acteurs. Les moyens identifiés pour cet argument visent donc à chercher à s'assurer que les personnes et les instances impliquées au sein de l'établissement d'enseignement et toute personne extérieure susceptible d'avoir accès à l'inférence d'évaluation connaissent les processus d'apprentissages et d'évaluation pour être plus à même d'évaluer elles-mêmes la crédibilité de l'inférence d'évaluation réalisée.

Un des premiers moyens pour chercher à s'assurer que l'inférence d'évaluation est acceptée des acteurs consiste en une assurance de qualité de la part des hautes instances de l'établissement d'enseignement. Le conseil d'administration du collège peut attester auprès des élèves, du public, du Ministère et de la Commission d'évaluation de l'enseignement collégial de la qualité, de l'équité et de l'équivalence des évaluations des apprentissages réalisées dans leur institution. Le fait qu'une haute instance atteste que des processus rigoureux d'évaluation ont été réalisés peut être suffisant pour en rassurer certains, mais un tel appel à l'autorité ne permet pas de faire connaître aux acteurs les processus impliqués.

Pour chercher à s'assurer que les processus d'apprentissage et d'évaluation sont connus des acteurs, la direction des études peut préciser les droits et responsabilités

de toute personne impliquée quant à l'évaluation des apprentissages dans la politique institutionnelle d'évaluation des apprentissages (PIEA) et veiller à la diffusion de cette dernière auprès de toute personne et instance impliquée à l'intérieur et à l'extérieur de l'établissement d'enseignement. La direction des études peut veiller à informer particulièrement les principaux concernés, les étudiants, sur les conditions de l'obtention du diplôme et sur les exigences du marché du travail. Elle peut aussi veiller à les informer sur les processus d'évaluation en diffusant un calendrier des évaluations pour l'ensemble du collège. Elle peut informer l'étudiant personnellement sur l'évolution de ses apprentissages en lui remettant un bulletin de mi-session et de fin de session.

Le département, via le plan de cours cadre, peut informer les professeurs sur les processus d'apprentissage impliqués par chaque cours, les approches pédagogiques et les modalités d'évaluation suggérées. En cherchant à ce que soient réglés les conflits en lien avec la politique institutionnelle d'évaluation des apprentissages (PIEA), le département peut chercher à s'assurer que les procédures d'évaluation sont comprises et acceptées, ce qui favorise par la même occasion l'acceptation de l'inférence d'évaluation.

Le professeur est l'acteur le plus près de l'étudiant et le plus à même de l'aider à comprendre les processus d'apprentissage et d'évaluation impliqués dans chaque cours et dans son programme d'étude. Il n'est donc pas étonnant de constater que les

moyens qui lui sont associés par rapport à cet argument sont nombreux. Le professeur peut informer directement l'étudiant sur le programme, sur les activités d'apprentissage et sur les modes d'évaluation des apprentissages via le plan de cours. Il peut aussi le faire à travers la planification des activités d'apprentissage et d'évaluation, ou à travers ses interventions individuelles auprès des étudiants, ses rétroactions. Ainsi, le professeur peut planifier les activités d'apprentissage et d'évaluation de façon à préparer progressivement l'étudiant à l'épreuve synthèse de programme et à l'atteinte des objectifs du cours et du programme. Il peut aussi planifier des activités d'évaluation formative qui, en plus de préparer l'étudiant à la prochaine évaluation, le renseigne sur son niveau de compétence, ce qui contribue à ce que l'étudiant ait une vision plus réaliste de ses compétences, et permet au professeur d'appuyer ses jugements d'évaluation quant au développement des compétences de l'étudiant au cours de la session. L'argument A5.2 renforce aussi l'importance accordée à la pratique d'informer l'étudiant sur les critères d'évaluation. Il est clair que si l'étudiant connaît les critères selon lesquels sa performance est évaluée, il est plus à même d'accepter comme valide l'inférence d'évaluation puisque celle-ci est clairement appuyée. En plus de la réponse aux critères, le professeur peut justifier son inférence d'évaluation à travers une rétroaction complète et détaillée à l'étudiant. Toutes ces pratiques diminuent les risques qu'il y ait dissonance cognitive entre la perception qu'a l'étudiant de ses propres compétences et l'inférence d'évaluation réalisée. L'étudiant qui a reçu suffisamment d'informations sur la progression de son apprentissage devrait être plus en mesure de juger de ses

performances et de sa démarche d'apprentissage. En ayant une perception réaliste de ses compétences, l'étudiant est plus susceptible d'accepter comme valide l'inférence d'évaluation, qui lui apparaît crédible parce qu'en accord avec sa propre perception.

L'étudiant a lui-même ses propres responsabilités pour s'assurer de connaître ses processus d'apprentissage et les processus d'évaluation. L'étudiant doit prendre connaissance de la politique institutionnelle d'évaluation des apprentissages (PIEA), des modalités et des règles d'évaluation propres au cours et au programme d'études, du plan de cours et du calendrier des évaluations. Il doit participer aux activités d'apprentissage ou veiller à faire les apprentissages attendus dans le cours. Pour l'étudiant, participer aux activités d'apprentissage en classe, étudier, faire des lectures, etc. contribuent à l'aider à mieux connaître ses propres processus d'apprentissage et à avoir une perception plus juste de ses propres compétences. S'il a une perception réaliste de ses propres compétences, l'étudiant sera plus à même d'accepter l'inférence d'évaluation. En somme, l'étudiant doit assurer le suivi de ses propres apprentissages, chercher à connaître ses processus d'apprentissage, utiliser les informations et les ressources dont il dispose pour ajuster sa démarche d'apprentissage et chercher l'aide appropriée au besoin. Ce sont des moyens que l'étudiant peut lui-même appliquer pour chercher à s'assurer de connaître les processus d'apprentissage et d'évaluation.

3.3.6 Répartition des moyens selon les niveaux d'inférence d'évaluation

Les moyens identifiés sont répartis selon les différents acteurs et instances de l'établissement d'enseignement. On constate que certains ont davantage de responsabilités que d'autres à travers les niveaux d'inférence du modèle. Plus on monte dans les niveaux d'inférence du modèle de Kane (2006) (I1 à I4), plus les responsabilités reviennent aux hautes instances du collège. On pouvait effectivement s'attendre à une telle répartition des responsabilités puisque les niveaux inférieurs du modèle, faisant référence à la planification du cours et des évaluation et à la notation, relèvent davantage des fonctions du professeur et du département, alors que les niveaux supérieurs du modèle, faisant référence à la définition du domaine et aux implications des inférences d'évaluation, relèvent davantage du département, du comité de programme, de la direction des études et du conseil d'administration du collège.

Concernant le niveau *Crédibilité* (C) qui a été ajouté au modèle initial de Kane (2006), très peu de moyens ont été identifiés pour l'argument *A5.1 L'ensemble des arguments est respecté de façon à ce que chacune des inférences considérées de façon individuelle soit convaincante*. À l'opposé, un très grand nombre de moyens ont été identifiés à l'argument *A5.2 Les processus d'apprentissage et d'évaluation sont connus des acteurs*.

Il est possible que peu de moyens spécifiques à l'argument A5.1 puissent être formulés parce que celui-ci se retrouve dans chacun des autres arguments et dans l'argument interprétatif dans son ensemble. Il s'agirait donc d'un niveau méta-évaluatif, d'un argument qui doit demeurer en tête à tous les niveaux de l'inférence d'évaluation.

Un très grand nombre de moyens ont été identifiés à l'argument A5.2 et la grande majorité de ces moyens sont spécifiques à cet argument, c'est-à-dire qu'ils ne se retrouvent dans aucun des arguments du modèle initial de Kane (2006). On peut toutefois qualifier cet argument de transversal en ce sens qu'il ne fait pas spécifiquement référence aux arguments du modèle de Kane (2006), mais qu'il est fonction de l'acceptation de ceux-ci. C'est aussi l'argument du modèle qui laisse la plus grande place à l'étudiant, alors que ce dernier n'a que peu ou pas d'incidence dans les autres niveaux du modèle.

CHAPITRE IV

DISCUSSION ET INTERPRÉTATIONS

Les analyses effectuées sur la base de la structure d'argumentation de validité de Kane (2006) ont révélé certaines lacunes en lien avec le contenu des politiques institutionnelles d'évaluation des apprentissages (PIEA) ou certaines incohérences relatives à la validité de l'inférence d'évaluation, particulièrement en ce qui a trait à la mesure, à l'évaluation des compétences auxiliaires et à l'évaluation des dispenses, des substitutions et des équivalences. Ces lacunes et incohérences sont exposées et discutées dans le présent chapitre qui se termine par une discussion et quelques pistes de réflexion sur la nature des moyens identifiés par l'analyse des politiques institutionnelles d'évaluation des apprentissages (PIEA).

4.1 Les lacunes au niveau de la mesure

Les politiques institutionnelles d'évaluation des apprentissages (PIEA) abordent l'évaluation sans aborder les questions relatives à la mesure. Elles font très peu référence au développement des procédures et des instruments de mesure. Ce faisant, elles négligent par la même occasion d'aborder les considérations liées aux erreurs systématiques pouvant être introduites dans l'inférence d'évaluation dès la construction d'un instrument ou d'une procédure de mesure.

Seules sont abordées des questions liées aux compétences linguistiques qui peuvent être des sources d'erreur systématique et elles sont rarement abordées sous cet angle. Kane (2006) aborde la question des compétences auxiliaires requises par les tests, lesquelles peuvent constituer des sources d'erreur systématique lors de la production de l'inférence d'évaluation. Selon lui, si le niveau de compétence dans une compétence auxiliaire, par exemple la lecture, requise par un test est bas comparativement au niveau de compétence de la population, la compétence auxiliaire en question ne devrait pas constituer une source de variance non pertinente pour la plupart des étudiants. Elle peut tout de même constituer une sérieuse source d'erreur systématique pour certains étudiants, par exemple ceux qui sont limités au niveau des habiletés de lecture. Par contre, si ces compétences sont assez variables dans la population testée pour influencer substantiellement les scores, elles peuvent être considérées comme des sources de variance non pertinente. Ces compétences auxiliaires deviennent alors des sources d'erreur systématique qu'il faut tenter de réduire au moment du développement des instruments et des procédures de mesure.

Certaines politiques institutionnelles d'évaluation des apprentissages (PIEA) suggèrent de procéder à la critique des outils et à la promotion des meilleurs instruments de mesure, mais elles ne mentionnent pas les critères à la base de cette critique ni ceux qui doivent servir à la sélection des meilleurs instruments. Les premiers tenants de l'approche par compétences ont souvent privilégié la validité de l'inférence d'extrapolation, c'est-à-dire la représentativité de l'évaluation par rapport

au domaine cible. Ainsi, on a vu se multiplier dans les classes les évaluations de types « performances » visant à rapprocher la tâche d'évaluation à la performance en contexte réel. Mais ce choix comporte un revers de médaille : celui de diminuer la validité de la généralisation. En effet, la tâche d'évaluation de type performance est conçue pour être plus représentative d'une tâche du domaine cible. Toutefois, puisque les évaluations de performances exigent beaucoup de temps, généralement elles incluent un nombre plus restreint de tâches. Dans de telles évaluations, puisque l'inférence de généralisation est effectuée à partir d'un plus petit échantillon de tâches, la validité de l'inférence de généralisation du score observé au score univers s'en trouve diminuée. Mais l'utilisation de tests hautement standardisés comporte aussi son revers de médaille. La validité de l'inférence de généralisation est certes augmentée par la représentativité de la mesure à l'univers de généralisation, mais la performance à un test standardisé s'éloigne beaucoup de la performance dans le domaine cible, en contexte réel. On se retrouve donc dans une impasse : renforcer l'extrapolation aux dépens de la généralisation en faisant en sorte que les tâches d'évaluation soient le plus possible représentatives du domaine cible, ou renforcer la généralisation aux dépens de l'extrapolation en utilisant un grand nombre de tâches hautement standardisées. (Kane, Cooks et Cohen, 1999 : voir Kane, 2006, p. 37)

Kane (2006, p. 37) suggère de développer des tests standardisés pour contrôler l'erreur aléatoire. Mais, comme la standardisation diminue la représentativité de l'univers de généralisation par rapport au domaine cible, il suggère de développer les tests de façon à toucher aux principales habiletés impliquées dans le domaine cible.

Ainsi, une preuve analytique de validité peut être développée lors du développement du test en faisant en sorte que le test soit le plus représentatif du domaine cible que possible. (Kane, 2006, p. 37)

Bien qu'une grande importance soit accordée aux activités de concertation dans le contenu des politiques institutionnelles d'évaluation des apprentissages analysées, il n'est presque jamais question des objectifs et du contenu de ces activités. La concertation des professeurs et des départements pourrait donc servir, entre autres, à identifier les principales habiletés du domaine cible qui devraient être évaluées (A3.1) et à identifier des sources d'erreurs systématiques et des moyens permettant de contrôler ces sources d'erreurs (A3.2) pour fournir aux professeurs des éléments qui puissent guider le développement des modalités d'évaluations et lorsque possible le développement de tests standardisés. Lors de l'élaboration des règles, le département peut aussi vérifier les aspects de la standardisation qui ont le plus de risques d'introduire une erreur dans l'interprétation des résultats d'évaluation. Le choix de la méthode d'évaluation (ex. : questions à réponses choisies, questions à développement, performances), des règles de notations (Clauser et *al.*, 2000 : voir Kane, 2006, p. 37), et des conditions d'observation (Brennan, 2001b : voir Kane, 2006, p. 37) peut introduire une variance non pertinente et par conséquent générer une erreur systématique. Kane (2006) suggère qu'au moins deux groupes développent les règles de notation de façon indépendante pour ensuite évaluer le degré d'accord entre les scores générés par les deux règles (Clauser, Harik et Clyman, 2000 : voir

Kane, 2006, p. 34). Il est parfois difficile de recourir à de telles méthodes pour établir indépendamment les règles de passation et de notation de chaque évaluation dans le contexte de l'enseignement. On peut tout de même songer à une méthode similaire pour la détermination des règles départementales d'évaluation des apprentissages en formant deux groupes lors des activités de concertation et en évaluant le degré d'accord entre les règles élaborées par les deux groupes. Ainsi, la détermination des règles départementales d'évaluation des apprentissages pourrait être le résultat de l'analyse des convergences et des divergences observées dans les règles élaborées par deux groupes et seraient ainsi basées sur l'expertise et le jugement des enseignants. Ces fonctions de la concertation ne sont pas mentionnées dans les politiques institutionnelles d'évaluation des apprentissages (PIEA) analysées.

À aucun endroit dans les politiques institutionnelles d'évaluation des apprentissages (PIEA) il n'est fait mention de l'entraînement des correcteurs. Il est bien fait mention de services d'encadrement et de perfectionnement des enseignants, mais aucune mention spécifique n'est faite en ce qui concerne la formation au sujet des biais de la passation et de la notation. Les correcteurs, enseignants ou étudiants employés, doivent être formés et sensibilisés aux sources de biais dans la passation et la notation pour être en mesure de les détecter et de minimiser leurs effets sur l'inférence de notation. Raïche (2008, p. 4) soutient l'idéal d'arriver à une correction plus objective des compétences, c'est-à-dire que les jugements d'appréciation puissent être remplacés « *par des actions de vérification de la présence ou de l'absence de*

manifestations de la compétence. » Ainsi, il encourage l'utilisation d'échelles descriptives analytiques ou globales plutôt que d'échelles d'appréciation dans l'évaluation des productions ou des performances d'un étudiant à l'enseignement supérieur. Encore une fois, la concertation des professeurs et des départements pourrait servir à identifier les habiletés principales du domaine cible qui devraient être évaluées, les critères et les standards d'évaluation.

Les politiques institutionnelles d'évaluation des apprentissages (PIEA) présentent donc des lacunes au niveau des aspects entourant la mesure. Celle-ci est d'importance capitale en évaluation des apprentissages et des considérations à son égard devraient figurer dans les politiques devant permettre d'orienter, d'encadrer et de soutenir les pratiques.

4.2 Le problème de l'intégration de compétences auxiliaires à l'évaluation

Les résultats de la présente recherche supportent les conclusions de Howe et Ménard (1993) selon lesquels les enseignants du collégial intègrent à leur mesure des cibles d'évaluation qui ne sont pas liées au trait, aux apprentissages ou à la compétence ciblée. Il s'agit de cibles d'évaluation telles que, comme le mentionnent Howe et Ménard (1993), le respect des exigences de présentation des travaux, la motivation et l'effort à l'apprentissage, la participation active de l'étudiant en classe

et la présence assidue aux cours. Les politiques institutionnelles d'évaluation des apprentissages (PIEA) mentionnent par exemple la présence aux cours ou aux activités d'évaluation, la présentation matérielle des travaux ou la qualité de la langue dont l'évaluation influence, souvent à la baisse, le résultat de l'évaluation.

L'évaluation des apprentissages actuellement réalisée au collégial, du moins tel qu'en témoignent les politiques institutionnelles d'évaluation des apprentissages (PIEA) analysées, ainsi qu'à d'autres ordres d'enseignement inclut l'évaluation de certains traits ou de certaines compétences auxiliaires sans toutefois qu'elles soient impliquées dans la mesure de la compétence. Le taux d'absentéisme en est un bon exemple. L'étudiant qui accumule un grand taux d'absentéisme accumule aussi les échecs. L'absence aux activités d'apprentissage et d'évaluation peut entraîner le retrait du droit de l'étudiant d'être évalué et conséquemment l'échec éventuel, et ce, quel que soit le niveau de compétence de l'étudiant dans le domaine cible. La mention échec est aussi attribuée lorsque l'abandon de cours n'a pas été effectué dans les délais prescrits. Or, les propriétés de la mention échec, si elles supportent les implications associées avec le trait ou la compétence mesurée, devraient signifier l'incompétence de l'étudiant par rapport au domaine cible. Cependant, ce n'est pas de la compétence de l'étudiant dans le domaine que témoigne l'échec dans ce cas précis, mais bien d'une compétence auxiliaire, soit la motivation ou tout autre trait ou compétence lié à l'absentéisme ou à l'abandon, qui seul détermine le score et biaise l'inférence d'évaluation. Cette erreur d'inférence peut nuire gravement à l'étudiant qui se voit

jugé sur son incompetence dans le domaine face à un échec qui cache en réalité une toute autre interprétation. Certains établissements d'enseignement utilisent la mention « EA » signifiant « Échec par abandon » pour palier ce problème d'interprétation du score. Au regard du modèle d'argumentation de la validité des inférences d'évaluation et des considérations éthiques, cette solution semble justifiée.

Les développements de la recherche en mesure et évaluation proposent aussi des façons d'intégrer simultanément à la mesure de la compétence une ou des mesures de la qualité des patrons de réponse de l'étudiant (Blais, Raïche et Magis, 2009; Raïche, Magis et Blais, 2008). Il s'agit de modélisations psychométriques qui permettent de corriger le score d'un étudiant malgré des tentatives de fraude pour augmenter son résultat, de l'inattention de sa part ou d'une fluctuation de son niveau de performance lors de la réalisation de la tâche d'évaluation. De cette façon, certaines sources de biais peuvent être contrôlées statistiquement et l'estimation du niveau de compétence de l'étudiant peut être plus représentatif des scores univers et cibles.

On peut aussi parfois remettre en question la pertinence de l'évaluation de la qualité de la langue en tant que compétence auxiliaire d'un domaine. En incluant la qualité de la langue dans l'évaluation, on affirme qu'il s'agit d'une compétence faisant partie du domaine et que son niveau de maîtrise est une condition au jugement de la compétence ou de l'incompétence d'un individu dans le domaine cible. On peut se demander s'il est socialement acceptable de conclure à l'incompétence d'une

personne dans un domaine, ou pire, de lui refuser l'accès à la formation dans ce domaine, en raison de son échec à l'évaluation d'une compétence auxiliaire qui, bien qu'étant liée au domaine cible, demeure auxiliaire. Les compétences en langue écrite et parlée sont des compétences auxiliaires évaluées dans la quasi-totalité des domaines d'éducation au Québec. Elles sont non seulement évaluées pendant la formation et sont une exigence à l'obtention du diplôme, mais elles sont également évaluées à l'entrée de la formation, en tant que critère d'admission. Cette évaluation est justifiée dans la plupart des cas par la nécessité de savoir bien s'exprimer à l'oral et à l'écrit une fois sur le marché du travail (sans bien sûr aborder la question de la préservation de la langue française au Québec, qui est une préoccupation sociale qui dépasse le thème de l'évaluation ici abordé). Toutefois, sur le marché du travail le niveau de maîtrise de la langue exigé diffère selon les domaines, tandis que cette différence n'est pas toujours reflétée dans la formation et l'évaluation. L'évaluation des compétences langagières devrait être adaptée au niveau exigé dans le domaine et, comme toute autre compétence incluse dans le domaine cible, elles devraient être définies par des experts du domaine. Une seule politique institutionnelle d'évaluation des apprentissages (PIEA) justifie l'évaluation de la langue en tant que compétence auxiliaire d'importance au sein du domaine cible. Dans certaines politiques institutionnelles d'évaluation des apprentissages (PIEA), il est attribué aux départements le rôle de spécifier les cas pour lesquels l'évaluation de la qualité de la langue est inapplicable. Il s'agit là d'une mention qui tient compte de la mise en

garde de Kane (2006) quant à l'évaluation des compétences auxiliaires et qui entre dans les exceptions à tenir compte dans l'inférence d'évaluation.

Des solutions, des moyens spécifiques doivent être explorés pour éviter que l'évaluation des compétences auxiliaires résulte en une erreur d'inférence.

4.3 Le problème de l'évaluation des dispenses, des substitutions et des équivalences

Le fait d'appliquer un modèle de validité à l'analyse des politiques institutionnelles d'évaluation des apprentissages (PIEA) a permis de cerner un problème récurrent en évaluation des apprentissages : celui de l'évaluation des dispenses, des substitutions et des équivalences.

Trois des cinq politiques institutionnelles d'évaluation des apprentissages (PIEA) analysées suggèrent de faire appel à l'accréditation de l'établissement comme garantie pour juger de la validité des inférences de notation et de généralisation lors de l'attribution des dispenses, des substitutions et des équivalences. Le service de l'établissement responsable de la reconnaissance des acquis doit dans ce cas supposer que les inférences de notation et de généralisation réalisées antérieurement, dans un cours un programme ou même dans un établissement différent, sont valides pour être en mesure de faire lui-même l'inférence d'extrapolation qui l'amène à reconnaître le score attribué au cours comme valide dans son interprétation au sein du domaine

cible. Pour ce faire, ce service n'a qu'une seule garantie de la validité des inférences de notation et d'extrapolation : « *que les objectifs du cours aient fait l'objet d'une évaluation formelle au sein d'un établissement reconnu* ». Les preuves sur laquelle s'appuie cette garantie sont : « *bulletin, attestation officielle de la réussite du ou des cours suivis, description des cours et plans de cours ou tout autre document pertinent* ». Ce qui se produit finalement, c'est que l'établissement introduit une valeur inconnue au sein de l'univers de généralisation qu'il suppose valide sur la seule base de l'accréditation de l'établissement d'enseignement qui fournit cette valeur et sur la similitude des objectifs et/ou des contenus planifiés.

Le problème rencontré avec cette façon de procéder est qu'il est impossible de vérifier que les arguments des inférences de notation et de généralisation sont respectés. Cela fait en sorte que l'argument *A5.1 L'ensemble des arguments est respecté de façon à ce que chacune des inférences considérées de façon individuelle soit convaincante* n'est pas respecté, ce qui invalide donc l'argument interprétatif dans son ensemble.

Deux des cinq politiques institutionnelles d'évaluation des apprentissages (PIEA) suggèrent d'élaborer et d'administrer un examen démontrant la maîtrise par l'étudiant de la compétence et des éléments de compétence prévus au cours faisant l'objet de la demande de reconnaissance. Il s'agit là d'une solution qui permet de chercher à s'assurer de la validité de l'inférence d'évaluation, puisque l'établissement reprend

alors le contrôle de chacun des niveaux d'inférence et peut ainsi chercher à s'assurer que chaque argument de validité est respecté, depuis la notation (I1) jusqu'à l'implication (I4), et chercher à s'assurer de la crédibilité de l'inférence d'évaluation réalisée.

4.4 La nature des moyens identifiés

L'analyse du contenu des politiques institutionnelles d'évaluation des apprentissages (PIEA) a permis d'identifier des moyens que peuvent mettre en place les institutions pour chercher à assurer la validité des inférences d'évaluation au regard des apprentissages des étudiants. Des moyens ont donc pu être identifiés pour chaque argument de la structure d'argumentation de Kane (2006) et même davantage en identifiant des moyens pour chercher à s'assurer de la crédibilité de l'inférence d'évaluation aux yeux des acteurs impliqués. Toutefois, les moyens identifiés dans le contenu des politiques institutionnelles d'évaluation des apprentissages (PIEA) sont de nature générale et correspondent davantage à des lignes directrices qu'à des moyens spécifiques.

En observant les moyennes et les écarts types du nombre d'énoncés classés dans chaque catégorie du modèle, on constate une grande variabilité entre les politiques institutionnelles d'évaluation des apprentissages (PIEA). Cet écart entre les politiques

institutionnelles d'évaluation des apprentissages (PIEA) pour l'importance accordée à chaque argument du modèle peut refléter les valeurs des collèges et ce qui les distingue au niveau de l'évaluation des apprentissages. L'objectif de cette recherche n'était pas de vérifier si les politiques institutionnelles d'évaluation des apprentissages (PIEA) contenaient tous les éléments permettant d'argumenter la validité des inférences d'évaluation, ni même d'effectuer une comparaison entre les politiques institutionnelles d'évaluation des apprentissages (PIEA). Toutefois, si le modèle suggéré et les hypothèses dégagées étaient subséquemment eux-mêmes validés par les acteurs du milieu collégial et par l'analyse des pratiques, le modèle et les lignes directrices identifiés pourraient éventuellement servir à l'évaluation et à l'élaboration des politiques institutionnelles d'évaluation des apprentissages (PIEA) dans une perspective d'argumentation de la validité des inférences d'évaluation des apprentissages.

Pour qu'elles puissent permettre aux établissements d'enseignement collégial *« d'orienter, d'encadrer et de soutenir les activités reliées à l'évaluation des apprentissages »* (Conseil des collèges, 1992, p. 38) et ainsi contribuer à assurer la validité des inférences d'évaluation des apprentissages, les politiques institutionnelles d'évaluation des apprentissages (PIEA) auraient avantage à inclure à leur contenu des moyens spécifiques d'application et d'évaluation des arguments de validité de la structure d'argumentation de la validité des inférences d'évaluation des apprentissages. Par exemple, « s'assurer de la concertation des enseignants sur les

objectifs, les contenus, les modes d'évaluation, les exigences et les seuils de réussite des cours et des programmes qui les concernent » est une ligne directrice qui devrait être élaborée en termes de moyens spécifiques permettant de centrer la concertation et de statuer sur chacun des éléments mentionnés et même davantage (sur le développement des modalités d'évaluation et des procédures de mesure par exemple comme il a déjà été mentionné à la section 4.1.1).

La recherche en mesure et évaluation en éducation met à la disposition des professionnels en éducation plusieurs développements qui constituent autant de moyens spécifiques qui peuvent être utilisés pour élaborer chacune des lignes directrices identifiées dans la présente recherche. Avec la structure d'argumentation de la validité des inférences d'évaluation des apprentissages et les lignes directrices identifiées dans la présente recherche, une analyse des pratiques effectives dans les établissements d'enseignement pourrait aussi permettre d'identifier des moyens d'application et d'évaluation spécifiques.

CONCLUSION

La présente recherche de nature exploratoire constitue une première tentative de validation de la structure d'argumentation interprétative de Kane (2006) par l'application de cette structure à l'analyse de politiques institutionnelles d'évaluation des apprentissages (PIEA) du réseau collégial québécois. Cette recherche poursuivait deux objectifs : 1) valider la structure d'argumentation interprétative de Kane (2006) par l'application de cette structure à l'analyse de politiques institutionnelles d'évaluation des apprentissages (PIEA) du réseau collégial québécois; 2) identifier des moyens que peuvent mettre en place les institutions pour chercher à assurer la validité des inférences d'évaluation au regard des apprentissages des étudiants.

La méthodologie retenue pour atteindre ces objectifs est l'analyse de contenu. Cinq politiques institutionnelles d'évaluation des apprentissages (PIEA), découpées en 424 énoncés distincts, ont été analysées au moyen d'une grille d'analyse de contenu construite à partir de la structure argumentaire développée par Kane (2006). Une modélisation schématique des moyens identifiés dans les politiques institutionnelles d'évaluation des apprentissages (PIEA) analysées a ensuite été réalisée, permettant

d'effectuer des regroupements, des subdivisions ou des ajouts à la structure d'argumentation interprétative de Kane (2006).

Les résultats obtenus ont permis de constater que le modèle de Kane (2006) peut effectivement s'appliquer au contexte de l'évaluation des apprentissages en enseignement collégial. De plus, les analyses effectuées ont permis de générer des hypothèses quant à l'exhaustivité, l'exclusivité et la pertinence des catégories du modèle de Kane (2006) dans ce contexte. Elles ont aussi permis de suggérer deux arguments supplémentaires à la structure d'argumentation initiale de façon à chercher à assurer la crédibilité de l'inférence d'évaluation aux yeux des acteurs impliqués.

Par la même occasion, cette recherche a permis l'identification de moyens, ou de lignes directrices, permettant de chercher à assurer l'argumentation de la validité des inférences d'évaluation en évaluation des apprentissages à l'enseignement supérieur. Des moyens spécifiques d'application et d'évaluation restent encore à être élaborés pour étoffer les lignes directrices identifiées dans cette recherche et pour ainsi faire en sorte que les politiques institutionnelles d'évaluation des apprentissages (PIEA) deviennent de véritables outils en évaluation des apprentissages, soutenant les pratiques et servant aux établissements d'enseignement, aux enseignants, aux professionnels et aux étudiants.

Étant donné le nombre limité de politiques institutionnelles d'évaluation des apprentissages (PIEA) analysées, on ne saurait prétendre à l'exhaustivité des moyens identifiés. Malgré cela, ceux-ci, constituant des lignes directrices pour chercher à assurer la validité des inférences d'évaluation des apprentissages, pourraient déjà servir de base de référence aux établissements d'enseignement collégial tout en laissant place aux développements pratiques et théoriques sur la validité des inférences d'évaluation en évaluation des apprentissages.

Des recherches ultérieures pourraient être réalisées pour identifier des moyens spécifiques d'application et d'évaluation de la structure d'argumentation de la validité des inférences d'évaluation des apprentissages proposée dans ce mémoire et adaptée de Kane (2006). Un processus d'analyse pourrait d'abord être réalisé pour identifier, pour chaque argument et pour chaque ligne directrice, les moyens suggérés et développés au sein des recherches récentes dans le domaine de la mesure et de l'évaluation en éducation. L'analyse de contenu de divers outils pédagogiques, tels que des plans de cours et des plans d'évaluation par exemple, ainsi que des entretiens auprès du personnel des établissements d'enseignement et auprès des étudiants pourraient ensuite être envisagés pour identifier les pratiques effectives dans les établissements d'enseignement. L'application des pratiques identifiées pourraient ensuite être évaluée à plus grande échelle au moyen d'une enquête par questionnaire réalisée dans l'ensemble des établissements d'enseignement collégial du Québec.

On pourrait alors considérer l'évaluation de l'efficacité perçue et/ou réelle de chaque pratique pour assurer la validité de l'inférence d'évaluation des apprentissages.

Étant donné l'envergure de la structure d'argumentation de la validité des inférences d'évaluation des apprentissages et la variété des pratiques pouvant y être associées, pour atteindre le niveau de spécificité nécessaire à l'identification de pratiques effectives, il semble nécessaire que chaque argument soit considéré de façon indépendante au sein d'un même programme de recherche portant sur la validité des inférences d'évaluation des apprentissages. La structure d'argumentation de la validité des inférences d'évaluation des apprentissages adaptée de Kane (2006) et les lignes directrices proposées dans ce mémoire, considérées dans un processus de recherche, pourraient être une première ébauche de modélisation à développer.

RÉFÉRENCES

- Auger, R. (2003, novembre). *Clarification conceptuelle et proposition d'opérationnalisation de quelques critères de scientificité de la recherche en éducation : le cas de la saturation et de la complétude*. Communication présentée à l'Association pour la recherche qualitative (ARQ), Trois-Rivières, Québec. Consulté le 13 mai 2007, tiré de <http://www.dep.uqam.ca/recherche/labform/ARQUQTR2003Criteresscientificite.pdf>.
- Auger, R. et Ségin, S. P. (1996). Validité globale d'une stratégie de testing adaptatif de maîtrise pour fins de certification scolaire au Québec. *Revue canadienne de l'éducation*, 21 (2), 143-154.
- Blais, J.-G., Raïche, G. et Magis, D. (2009). La détection des patrons de réponses problématiques dans le contexte des tests informatisés. Dans J.-G. Blais (Dir.) : *Évaluation des apprentissages et technologie de l'information et de la communication. Enjeux, applications et modèles de mesure*. Ste-Foy, Québec : Presses de l'Université Laval.
- Brennan, R. L. (2006). Perspectives on the evolution and future of educational measurement. Dans R. L. Brennan (Dir.) : *Educational measurement* (4^e édition). Westport, Connecticut : Praeger Publishers.
- Commission d'évaluation de l'enseignement collégial (1994a). *La Commission d'évaluation de l'enseignement collégial : sa mission et ses orientations*. Québec, Québec : Ministère de l'Éducation du Québec.
- Commission d'évaluation de l'enseignement collégial (1994b). *L'évaluation des politiques institutionnelles d'évaluation des apprentissages – Cadre de référence*. Québec, Québec : Ministère de l'Éducation du Québec.
- Conseil des collèges (1992). *L'enseignement collégial : des priorités pour un renouveau de la formation*. Québec, Québec : Gouvernement du Québec.
- Cronbach, L. J. (1971). Test validation. Dans R. L. Thorndike (Dir.), *Educational Measurement* (4^e édition). Washington, District of Colombia : American council on education.

- Cronbach, L. J. (1980). Validity on parole : How can we go straight. Dans W. B. Schrader (Dir.), *New directions for testing and measurement : Measuring achievement : progress over a decade*. San Fransisco, California : Jossey-Bass.
- Cronbach, L. J. (1988). Five perspectives on validity argument. Dans H. Wainer et H. Braun (Dirs.), *Test validity*. Hillsdale, New Jersey : Lawrence Erlbaum associates.
- Cronbach, L. J., Linn, R. L., Brennan, R. L. et Haertel, E. H. (1997). Generalizability analysis for performance assessments of student achievement or school effectiveness. *Educational and Psychological Measurement*, 57 (3), 373-399.
- Cureton, E. E. (1951). *Validity*. Dans E. F. Lindquist (Dir.), *Educational Measurement*. Washington, District of Colombia : American council on education.
- De Ketele, J.-M. et F.-M. Gerard (2005). La validation des épreuves d'évaluation selon l'approche par les compétences. *Mesure et évaluation en éducation*, 28 (3), 1-26.
- Gipps, C. V. (1994). *Beyond testing : towards a theory of educational assessment*. Washington, District of Colombia : Falmer Press.
- Howe, R. et L. Ménard (1993). *Croyances et pratiques en évaluation des apprentissages*. Rapport de recherche. Laval, Québec : Cégep Montmorency.
- Howe, R. et L. Ménard (1994). Croyances et pratiques en évaluation des apprentissages. *Pédagogie collégiale*, 7 (3), 21-27.
- Kane, M.T. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112, 527-535.
- Kane, M. (2004). Certification testing as an illustration of argument-based validation. *Measurement: Interdisciplinary Research and Perspectives*, 2, 135-170.
- Kane, M. T. (2006). Validation. Dans R. L. Brennan (Dir.), *Educational measurement* (4^e édition). Westport, Connecticut : Praeger Publishers.
- L'Écuyer, R. (1987). L'analyse de contenu : notion et étapes. Dans J.-P. Deslauriers (Dir.) : *Les méthodes de recherche qualitatives*. Sillery, Québec : Presses de l'Université du Québec.

- Lessard-Hébert, M., Goyette, G. et Boutin, G. (1996). *La recherche qualitative, fondements et pratiques*. Montréal, Québec : Éditions Nouvelles.
- LICEF, Télé-université. [En ligne]. Site web disponible le 25 mai 2009 : <http://www.liceftel.uqam.ca>.
- Lissitz, R. W. (Dir.) (2009). *The concept of validity: revisions, new directions, and applications*. Charlotte, North Carolina : Information Age Publishing.
- Messick, S. (1989). Validity. Dans R. L. Linn (Dir.), *Educational measurement* (3^e édition). New-York, New-Jersey : American Council on Education et Macmillan.
- Ministère de l'Éducation, des Loisirs et du Sport (2007). *Tableau 07 L'effectif scolaire des établissements d'enseignement collégial selon la région administrative, l'établissement, le type de formation, le service d'enseignement et le réseau d'enseignement (2004 ou 2005 ou 2006). Statistiques détaillées sur l'éducation*. [En ligne]. Disponible le 8 avril 2008 : http://www.mels.gouv.qc.ca/stat/Stat_det/coll_eff.htm.
- Ministère de l'Enseignement supérieur et de la Science (1993). *Des collèges pour le Québec du XXI^e siècle*. Québec, Québec : Gouvernement du Québec.
- Mislevy, R. J. (1994). Evidence and inference in educational assessment. *Psychometrika*, 59, 439-483.
- Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2004). *A brief introduction to evidence-centered design*. CSE Report 632. Los Angeles, Californie : The National Center for Research on Evaluation, Standards, Student Testing (CRESST), Center for the Study of Evaluation (CSE), University of California. [En ligne]. Disponible le 13 mars 2010 : <http://www.cse.ucla.edu/products/reports/r632.pdf>.
- Mislevy, R., & Haertel, G. (2006a). *Implications of Evidence-Centered Design for Educational Testing*. Draft PADI Technical Report 17. Menlo Park, Californie : SRI International. [En ligne]. Disponible le 13 mars 2010 : http://padi.sri.com/downloads/TR17_EMIP.pdf.
- Mislevy, R. J., & Haertel, G. D. (2006b). Implications for evidence-centered design for educational testing. *Educational Measurement : Issues and Practice*, 25 (4), 6-20.

- Mislevy, R. J., Steinberg, L. S., Almond, R. G. et Lukas, J. F. (2006). Concepts, terminology, and basic models of evidence-centered design. Dans D. M. Williamson, R. J. Mislevy, et I. I. Bejar (Eds.), *Automated scoring of complex tasks in computer-based testing*. Mahwah, New Jersey : Lawrence Erlbaum Associates.
- Muchielli, R. (1971). *L'examen psychotechnique. Applications pratiques*. Paris, France : Les Éditions ESF.
- Paquette-Côté, K. (2009). *L'utilisation des politiques institutionnelles d'évaluation des apprentissages pour assurer la validité des inférences d'évaluation des apprentissages*. 31e session d'étude de l'Association pour le développement des méthodologies d'évaluation en éducation (ADMÉE-Canada), Québec.
- Paquette-Côté, K. et Raïche, G. (2008). *L'argumentation de la validité des inférences d'évaluation dans les politiques institutionnelles d'évaluation des apprentissages des établissements d'enseignement collégial québécois*. 30e session d'étude de l'Association pour le développement des méthodologies d'évaluation en éducation (ADMÉE-Canada), Trois-Rivières, Québec.
- Paquette-Côté, K. et Raïche, G. (2009). *Analyse de l'argumentation de la validité des inférences d'évaluation dans les politiques institutionnelles d'évaluation des apprentissages des établissements d'enseignement collégial québécois*. Communication présentée au 77e Congrès de l'Acfas, Université d'Ottawa, Ottawa, Ontario.
- Paquette-Côté, K. et Raïche, G. (à paraître). La validité des inférences d'évaluation dans les politiques institutionnelles d'évaluation des apprentissages des établissements d'enseignement collégial québécois. Dans Raïche, G. et Paquette-Côté, K. (Dir.), *Des mécanismes pour assurer la validité de l'interprétation de la mesure en éducation*. Ste-Foy, Québec : Presses de l'Université du Québec.
- Raïche, G. (2008). *L'évaluation des apprentissages à l'enseignement supérieur : vers une vision intégrative de l'évaluation des apprentissages* (2e édition). Montréal, Québec : Université du Québec à Montréal.
- Raïche, G., Magis, D. et Blais, J.-G. (2008). *Multidimensional item response theory models integrating additional inattention, pseudo-guessing, and discrimination person parameters*. Communication présentée lors du congrès annuel de la Psychometric Society, Durham, New Hampshire.

- Rivard, J. (2007). *Logiciel MotPlus. Éditeur de modèles de connaissances. Manuel de l'utilisateur. Version du logiciel : 1.6.5.* Québec, Québec : LICEF, Télé-Université.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences.* Saint-Laurent, Québec : Éditions du Renouveau pédagogique Inc.
- Toulmin, S. E (1964). *The uses of argument.* Cambridge, United Kingdom : Cambridge University Press.
- Van der Maren, J.-M. (1995). *Méthodes de recherche pour l'éducation.* Montréal, Québec : Les Presses de l'Université de Montréal.
- Wiggins, G. P. (1993). *Assessing student performance : exploring the purpose and limits of testing.* San Francisco, California : Jossey-Bass.

APPENDICE A

GRILLE D'ANALYSE DU CONTENU MANIFESTE DES POLITIQUES INSTITUTIONNELLES D'ÉVALUATION DES APPRENTISSAGES

Grille d'analyse du contenu manifeste des politiques institutionnelles d'évaluation des apprentissages (PIEA) selon le modèle de validation de l'argument interprétatif pour l'interprétation d'un trait de Kane (2006)

Description de la grille d'analyse de contenu :

La première colonne permet d'identifier la politique institutionnelle d'évaluation des apprentissages (PIEA) analysée.

La deuxième colonne contient l'unité d'analyse, soit un énoncé correspondant à une phrase ou un ensemble de phrases exprimant une seule idée.

Dans la troisième colonne est identifié pour chaque extrait l'argument de validité du modèle de Kane (2006) auquel il fait référence. Dans cette colonne, l'unité d'analyse peut être codée par un seul ou par plusieurs arguments de validité du modèle de Kane (2006). L'unité d'analyse peut aussi être codée « Autre » lorsque l'énoncé se rapporte à la validité de l'inférence d'évaluation, mais qu'il ne correspond à aucun argument du modèle de Kane (2006) en particulier.

La quatrième colonne permet de coder si un moyen d'application et/ou d'évaluation de l'argument est identifié dans l'énoncé.

La cinquième colonne est réservée aux interprétations de l'analyste pouvant être, selon le cas et sans s'y limiter : le moyen d'application et/ou d'évaluation de l'argument, une interprétation au sujet de l'énoncé ou de l'argument de validité auquel il fait référence, une interprétation au sujet du modèle de Kane (2006).

[illegible]

APPENDICE B

MOYENS POUR CHERCHER À ASSURER LA VALIDITÉ DES INFÉRENCES D'ÉVALUATION DES APPRENTISSAGES

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A1.1 Les règles de passation et de notation sont appropriées

A1.1 Les règles de passation et de notation sont appropriées.	
Professeur	<ol style="list-style-type: none"> 1. S'impliquer et participer et aux activités de concertation de programme. 2. Planifier les activités d'apprentissage et d'évaluation en conformité avec les plans cadres ou avec les objectifs et standards de chacun des cours. <ol style="list-style-type: none"> 2.1 Rédiger un plan de cours conforme à la politique institutionnelle d'évaluation des apprentissages (PIEA), aux règles départementales d'évaluation des apprentissages, au programme et au plan de cours cadre. <ol style="list-style-type: none"> 2.1.1 Indiquer dans le plan de cours les activités d'évaluation sommative des apprentissages et leurs modalités : la nature des épreuves (travaux, examens ou autres) servant à l'établissement de la note finale, leur pondération, leur durée et la date prévue. 2.1.2 Soumettre le plan de cours au département aux fins d'approbation. 2.2 Accomplir les tâches d'évaluation : fixer des objectifs, des standards et des critères d'évaluation; choisir et élaborer des instruments d'évaluation des apprentissages des étudiants. <ol style="list-style-type: none"> 2.2.1 Établir clairement les critères et les standards d'évaluation. 2.2.2 Développer, réviser et adapter les procédures et les instruments de mesure et d'évaluation. 2.2.3 À la demande de la direction des études, fournir le corrigé de toute évaluation.
Département	<ol style="list-style-type: none"> 1. S'assurer de la concertation des enseignants sur les objectifs, les contenus, les modes d'évaluation, les exigences et les seuils de réussite des cours et des programmes qui les concernent. <ol style="list-style-type: none"> 1.1 Rechercher l'équivalence des objectifs d'apprentissage, des procédés et des critères d'évaluation dans un même cours dispensé par plusieurs enseignants. <ol style="list-style-type: none"> 1.1.1 S'assurer que le même plan de cours soit utilisé pour un cours donné à plusieurs groupes d'étudiants, à la même session dans un même secteur. 1.2 Rechercher l'équivalence des évaluations par la concertation entre les professeurs d'un même département et entre les départements au sujet des modes et des instruments d'évaluation et à propos du degré d'exigence et des seuils de réussite. <ol style="list-style-type: none"> 1.2.1 S'assurer que soit élaboré un examen commun pour l'évaluation finale un cours donné à plusieurs groupes d'étudiants, à la même session dans un même secteur.

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A1.1 Les règles de passation et de notation sont appropriées

Département	<p>2. Établir, transmettre et faire connaître des règles départementales d'évaluation des apprentissages conformes à la PIEA et veiller à leur application.</p> <ul style="list-style-type: none"> ▪ <i>Exemple : établissement d'une règle selon laquelle seul ce qui est prévu au plan de cours et ce qui a été enseigné doit faire l'objet d'une évaluation.</i> <p>2.1 Élaborer, adopter, diffuser des règles d'évaluation propres au programme et veiller à leur application.</p> <p>2.2 Établir les modalités d'évaluation spécifiques et préciser les standards de réussite pour chacun des cours dont il est responsable.</p> <ul style="list-style-type: none"> ▪ <i>Exemple : la note de passage est de 60 %. Cette note traduit l'atteinte minimale des objectifs et des standards visés par le cours.</i> <p>2.3 Déterminer les modalités spécifiques de présentation des travaux et les exigences propres aux laboratoires et aux stages.</p> <p>2.4 S'assurer que les activités et les instruments d'évaluation respectent les principes de la politique institutionnelle d'évaluation des apprentissages (PIEA), les modalités départementales d'évaluation des apprentissages et le plan cadre.</p>
	<p>3. Élaborer un plan de cours cadre pour chacun des cours dont il est responsable.</p> <p>3.1 Transmettre les plans cadres aux professeurs.</p> <p>3.2 Établir et faire approuver par la direction des études les modalités d'analyse et d'approbation des plans de cours dont il est responsable.</p> <ul style="list-style-type: none"> 3.2.1 Approuver les plans de cours. 3.2.2 Approuver les modifications des plans de cours en cours de session. 3.2.3 Transmettre les plans de cours à la direction des études en faisant état de leur approbation.
	<p>4. Se renseigner sur les besoins de perfectionnement ou de soutien en matière d'évaluation des apprentissages et planifier lui-même des activités de perfectionnement.</p> <p>4.1 Communiquer les besoins de perfectionnement ou de soutien en matière d'évaluation des apprentissages à la direction des études.</p>
Service de formation continue	1. Offrir des services d'encadrement et de perfectionnement au personnel enseignant pour les soutenir dans leurs pratiques d'évaluation des apprentissages.
Comité de programme	1. Déterminer et mettre en application des procédures d'évaluation conformes à la politique institutionnelle d'évaluation des apprentissages (PIEA) et aux règles départementales d'évaluation des apprentissages.

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages

A1.1 Les règles de passation et de notation sont appropriées

Direction des études	1. Consulter la commission des études au sujet des règles d'évaluation propres aux programmes et de leur approbation.
	2. Décrire et diffuser les procédures administratives requises pour la passation et la notation des activités d'évaluation.
	3. Lors de l'attribution d'une dispense, d'une équivalence ou d'une substitution, élaborer et administrer un examen démontrant la maîtrise par l'étudiant de la compétence et des éléments de compétence prévus au cours faisant l'objet de la demande.
	4. S'assurer que soient organisées des activités de perfectionnement en matière d'évaluation des apprentissages.
	5. Approuver les règles départementales d'évaluation des apprentissages.
	6. Approuver les modalités d'analyse et d'approbation des plans de cours. 6.1 Conserver et rendre accessible pour consultation les plans de cours approuvés.
	7. Lorsque nécessaire, recueillir et vérifier le corrigé des évaluations.

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A1.2 Les règles de passation et de notation sont appliquées telles que spécifiées

A1.2 Les règles de passation et de notation sont appliquées telles que spécifiées.	
Professeur	1. S'impliquer et participer et aux activités de concertation de programme. 2. Accomplir les tâches d'évaluation : fixer des objectifs, des standards et des critères d'évaluation; choisir et élaborer des instruments d'évaluation des apprentissages des étudiants. 2.1 Connaître et appliquer la politique institutionnelle d'évaluation des apprentissages (PIEA) et les règles départementales d'évaluation des apprentissages lors de la passation et la notation des activités d'évaluation. 2.2 Assumer la responsabilité de la passation des activités d'évaluation. 3. Fonder l'évaluation sur l'atteinte des objectifs d'apprentissage annoncés dans le plan de cours ou sur l'atteinte des éléments de la compétence, selon les standards de réussite prescrits ou définis par le département.
Département	1. S'assurer de la concertation des enseignants sur les objectifs, les contenus, les modes d'évaluation, les exigences et les seuils de réussite des cours et des programmes qui les concernent. 2. Transmettre et faire connaître des règles départementales d'évaluation des apprentissages conformes à la PIEA et veiller à leur application. 3. Élaborer, adopter, diffuser des règles d'évaluation propres au programme et veiller à leur application. 4. Se renseigner sur les besoins de perfectionnement ou de soutien en matière d'évaluation des apprentissages et planifier lui-même des activités de perfectionnement. 4.1 Communiquer les besoins de perfectionnement ou de soutien en matière d'évaluation des apprentissages à la direction des études.
Service de formation continue	1. Offrir des services d'encadrement et de perfectionnement au personnel enseignant pour les soutenir dans leurs pratiques d'évaluation des apprentissages.
Direction des études	1. Décrire et diffuser les procédures administratives requises pour la passation et la notation des activités d'évaluation. 2. S'assurer que soient organisées des activités de perfectionnement en matière d'évaluation des apprentissages. 3. Assurer l'administration de toute épreuve uniforme imposée par le ministre. 4. Lors de l'attribution d'une dispense, d'une équivalence ou d'une substitution, élaborer et administrer un examen démontrant la maîtrise par l'étudiant de la compétence et des éléments de compétence prévus au cours faisant l'objet de la demande. 5. Soumettre un rapport écrit sur le processus d'évaluation des apprentissages au conseil d'administration à la fin de chaque session.

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A1.3 La passation et la notation sont exemptes de biais

A1.3 La passation et la notation sont exemptes de biais.	
Étudiant	1. Connaître et appliquer la politique institutionnelle d'évaluation des apprentissages (PIEA) et les règles départementales d'évaluation des apprentissages lors de la passation des activités d'évaluation.
Professeur	1. Connaître et appliquer la politique institutionnelle d'évaluation des apprentissages (PIEA) et les règles départementales d'évaluation des apprentissages lors de la passation et la notation des activités d'évaluation.
	2. Établir clairement les critères et les standards d'évaluation.
	3. Répartir l'évaluation sommative sur plusieurs activités d'évaluation en cours de session.
	4. Fournir aux étudiants les consignes de présentation des travaux.
	5. Assumer la responsabilité de la passation des activités d'évaluation.
	6. Prendre les moyens nécessaires pour garantir la confidentialité des instruments d'évaluation avant leur administration aux étudiants.
	7. S'assurer que la notation porte sur l'évaluation des apprentissages individuels. ▪ Exemple : la notation des apprentissages individuels dans les travaux d'équipe ne devrait pas être biaisée par la performance des autres membres de l'équipe.
	8. Adapter les activités d'évaluation aux besoins particuliers de certains groupes ou certains individus pour lesquels les activités d'évaluations planifiés pourraient constituer une source de biais systématique.
Département	1. S'assurer de la concertation des enseignants sur les objectifs, les contenus, les modes d'évaluation, les exigences et les seuils de réussite des cours et des programmes qui les concernent. 1.1 Rechercher l'équivalence des objectifs d'apprentissage, des procédés et des critères d'évaluation dans un même cours dispensé par plusieurs enseignants. 1.1.1 S'assurer que le même plan de cours soit utilisé pour un cours donné à plusieurs groupes d'étudiants, à la même session dans un même secteur. 1.2 Rechercher l'équivalence des évaluations par la concertation entre les professeurs d'un même département et entre les départements au sujet des modes et des instruments d'évaluation et à propos du degré d'exigence et des seuils de réussite. 1.2.1 S'assurer que soit élaboré un examen commun pour l'évaluation finale un cours donné à plusieurs groupes d'étudiants, à la même session dans un même secteur.
	2. Se renseigner sur les besoins de perfectionnement ou de soutien en matière d'évaluation des apprentissages et planifier lui-même des activités de perfectionnement.

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A1.3 La passation et la notation sont exemptes de biais

Département	<p>2.1 Communiquer les besoins de perfectionnement ou de soutien en matière d'évaluation des apprentissages à la direction des études.</p> <p>3. Établir, transmettre et faire connaître des règles départementales d'évaluation des apprentissages conformes à la politique institutionnelle d'évaluation des apprentissages (PIEA) et veiller à leur application</p> <ul style="list-style-type: none"> ▪ Exemple : établissement d'une règle selon laquelle un étudiant ne peut être soumis à plus de deux activités d'évaluation de 10 % et plus, d'une durée de 50 minutes et plus dans une même journée. ▪ Exemple : établissement d'une règle selon laquelle la pondération de l'examen final ne doit pas excéder 25% de la note totale d'un cours.
Service de formation continue	1. Offrir des services d'encadrement et de perfectionnement au personnel enseignant pour les soutenir dans leurs pratiques d'évaluation des apprentissages.
Comité de programme	1. S'assurer que les moyens d'évaluation spécifiés dans les plans de cours sont équitables.
Direction des études	<p>1. Décrire et diffuser les procédures administratives requises pour la passation et la notation des activités d'évaluation.</p> <p>2. S'assurer que soient organisées des activités de perfectionnement en matière d'évaluation des apprentissages.</p> <p>3. Assurer l'administration de toute épreuve uniforme imposée par le ministre.</p> <p>4. Lors de l'attribution d'une dispense, d'une équivalence ou d'une substitution, élaborer et administrer un examen démontrant la maîtrise par l'étudiant de la compétence et des éléments de compétence prévus au cours faisant l'objet de la demande.</p> <p>5. Structurer le calendrier des évaluations pour l'ensemble du collège de façon à chercher à s'assurer qu'aucun étudiant n'ait plus de deux évaluations par jour.</p> <p>6. S'assurer que les conditions environnementales d'évaluation sont adéquates (ex. : grandeur et ameublement des locaux, lumière, contrôle des bruits environnants).</p> <p>7. Minimiser les possibilités de plagiat et de tricherie lors de la passation des activités d'évaluation.</p> <p>8. Assurer la sécurité des examens et des cahiers de réponse.</p>

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A1.4 Les données s'adaptent à tous les modèles d'échelles de mesure employés dans la notation

A1.4 Les données s'adaptent à tous les modèles d'échelles de mesure employés dans la notation.	
Professeur	1. Accomplir les tâches d'évaluation : fixer des objectifs, des standards et des critères d'évaluation; choisir et élaborer des instruments d'évaluation des apprentissages des étudiants. 1.1 Développer, réviser et adapter les procédures et les instruments de mesure et d'évaluation. 1.2 Établir clairement l'échelle de notation employée et la façon dont les cotes sont interprétées numériquement.
Département	1. S'assurer de la concertation des enseignants sur les objectifs, les contenus, les modes d'évaluation, les exigences et les seuils de réussite des cours et des programmes qui les concernent. 1.1 Rechercher l'équivalence des évaluations par la concertation entre les professeurs d'un même département et entre les départements au sujet des modes et des instruments d'évaluation et à propos du degré d'exigence et des seuils de réussite. 2. Se renseigner sur les besoins de perfectionnement ou de soutien en matière d'évaluation des apprentissages et planifier lui-même des activités de perfectionnement. 2.1 Communiquer les besoins de perfectionnement ou de soutien en matière d'évaluation des apprentissages à la direction des études.
Service de formation continue	1. Offrir des services d'encadrement et de perfectionnement au personnel enseignant pour les soutenir dans leurs pratiques d'évaluation des apprentissages.
Direction des études	1. S'assurer que soient organisées des activités de perfectionnement en matière d'évaluation des apprentissages.

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A2.1 L'échantillon des observations est représentatif de l'univers de généralisation

A2.1 L'échantillon des observations est représentatif de l'univers de généralisation.	
Professeur	<p>1. Planifier les activités d'apprentissage et d'évaluation en conformité avec les plans cadres ou avec les objectifs et standards de chacun des cours.</p> <p>1.1 Rédiger un plan de cours conforme à la politique institutionnelle d'évaluation des apprentissages (PIEA), aux règles départementales d'évaluation des apprentissages, au programme et au plan de cours cadre.</p> <p>1.1.1 Indiquer dans le plan de cours la compétence visée par le cours et l'apport du cours à la réalisation des objectifs du programme.</p> <p>1.1.2 Indiquer dans le plan de cours les objectifs de formation du cours présentés en énoncés de compétence et en standards.</p> <p>1.1.3 Indiquer dans le plan de cours les objectifs du cours ; les contenus couverts par le cours ; les activités d'enseignement et d'apprentissage.</p> <p>1.1.4 Indiquer dans le plan de cours les activités d'évaluation sommative des apprentissages et leurs modalités : la nature des épreuves (travaux, examens ou autres) servant à l'établissement de la note finale, leur pondération, leur durée et la date prévue.</p> <p>1.1.5 Soumettre le plan de cours au département aux fins d'approbation.</p> <p>1.2 Accomplir les tâches d'évaluation : fixer des objectifs, des standards et des critères d'évaluation; choisir et élaborer des instruments d'évaluation des apprentissages des étudiants.</p> <p>1.2.1 Concevoir des activités d'évaluation des apprentissages conformes aux objectifs du cours ou aux compétences à développer.</p> <p>1.2.2 Répartir l'évaluation sommative sur plusieurs activités d'évaluation en cours de session.</p>
Département	<p>1. S'assurer de la concertation des enseignants sur les objectifs, les contenus, les modes d'évaluation, les exigences et les seuils de réussite des cours et des programmes qui les concernent.</p> <p>1.1 Rechercher l'équivalence des objectifs d'apprentissage, des procédés et des critères d'évaluation dans un même cours dispensé par plusieurs enseignants.</p> <p>1.1.1 S'assurer que le même plan de cours soit utilisé pour un cours donné à plusieurs groupes d'étudiants, à la même session dans un même secteur.</p>

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages

A2.1 L'échantillon des observations est représentatif de l'univers de généralisation

Département	<p>1.2 Rechercher l'équivalence des évaluations par la concertation entre les professeurs d'un même département et entre les départements au sujet des modes et des instruments d'évaluation et à propos du degré d'exigence et des seuils de réussite.</p> <p>1.2.1 S'assurer que soit élaboré un examen commun pour l'évaluation finale un cours donné à plusieurs groupes d'étudiants, à la même session dans un même secteur.</p> <p>2. Élaborer des règles d'évaluation propres au programme qui précisent comment la maîtrise de la compétence est évaluée pour obtenir la note de passage.</p> <p>2.1 Pondérer l'évaluation des apprentissages en fonction de la valeur relative accordée aux objectifs d'apprentissage et aux compétences dans chaque cours dont il est responsable.</p> <p>2.2 Déterminer les modalités spécifiques de présentation des travaux et les exigences propres aux laboratoires et aux stages.</p> <p>3. Élaborer un plan de cours cadre pour chacun des cours dont il est responsable.</p> <p><i>Le plan cadre présente les objectifs d'apprentissage et les contenus essentiels du cours.</i></p> <p>3.1 Transmettre les plans cadres aux professeurs.</p> <p>3.2 Établir et faire approuver par la direction des études les modalités d'analyse et d'approbation des plans de cours dont il est responsable.</p> <p>3.2.1 Approuver les plans de cours.</p> <p>3.2.2 Approuver les modifications des plans de cours en cours de session.</p> <p>3.2.3 Transmettre les plans de cours à la direction des études en faisant état de leur approbation.</p> <p>4. S'assurer que les activités et les instruments d'évaluation respectent les principes de la politique institutionnelle d'évaluation des apprentissages (PIEA), les modalités départementales d'évaluation des apprentissages et le plan cadre.</p> <p>5. Valider, administrer et évaluer l'épreuve synthèse de programme.</p> <p>6. Se renseigner sur les besoins de perfectionnement ou de soutien en matière d'évaluation des apprentissages et planifier lui-même des activités de perfectionnement.</p> <p>6.1 Communiquer les besoins de perfectionnement ou de soutien en matière d'évaluation des apprentissages à la Direction des études.</p>
Service de formation continue	<p>1. Offrir des services d'encadrement et de perfectionnement au personnel enseignant pour les soutenir dans leurs pratiques d'évaluation des apprentissages.</p>

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A2.1 L'échantillon des observations est représentatif de l'univers de généralisation

Comité de programme	<p>1. Établir le contenu des cours en lien avec les objectifs et les standards du programme, du domaine et des exigences du collège et du Ministère.</p> <p>2. S'assurer que les instruments et les méthodes d'évaluation de tous les cours sont conformes à la politique institutionnelle d'évaluation des apprentissages (PIEA), aux règles départementales d'évaluation des apprentissages, au programme et aux plans de cours cadre.</p> <p>2.1 S'assurer que les moyens d'évaluation spécifiés dans les plans de cours sont appropriés par rapport aux objectifs et aux standards du cours et du programme.</p> <p>3. Assumer la responsabilité de l'élaboration de l'épreuve synthèse de programme et participer à son administration.</p> <p>▪ <i>Exemple : L'épreuve synthèse tient compte des objectifs et des standards déterminés dans le programme ministériel, du profil de sortie et de la perspective d'équivalence intra-institutionnelle de l'évaluation et de la synthèse.</i></p>
Direction des études	<p>1. S'assurer que soient organisées des activités de perfectionnement en matière d'évaluation des apprentissages.</p> <p>2. Approuver les modalités d'analyse et d'approbation des plans de cours.</p> <p>2.1 S'assurer de la conformité des plans de cours avec les objectifs, les standards et, s'il y a lieu, les activités d'apprentissage établis pour chacun des cours.</p> <p>2.2 Conserver et rendre accessible pour consultation les plans de cours approuvés.</p> <p>3. Approuver les modalités générales de l'épreuve synthèse de chaque programme conduisant au diplôme d'études collégiales.</p> <p>4. Lors de l'attribution d'une dispense, d'une équivalence ou d'une substitution, élaborer et administrer un examen démontrant la maîtrise par l'étudiant de la compétence et des éléments de compétence prévus au cours faisant l'objet de la demande.</p>
Commission des études	<p>1. Donner son avis à la direction des études sur les modalités générales de l'épreuve synthèse de programme et sur l'harmonisation des pratiques d'évaluation dans l'ensemble des programmes.</p>

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A2.2 L'échantillon des observations est suffisamment grand pour contrôler l'erreur aléatoire

A2.2 L'échantillon des observations est suffisamment grand pour contrôler l'erreur aléatoire.	
Professeur	<p>1. Rédiger un plan de cours conforme à la politique institutionnelle d'évaluation des apprentissages (PIEA), aux règles départementales d'évaluation des apprentissages, au programme et au plan de cours cadre.</p> <p>1.1 Indiquer dans le plan de cours les activités d'évaluation sommative des apprentissages et leurs modalités : la nature des épreuves (travaux, examens ou autres) servant à l'établissement de la note finale, leur pondération, leur durée et la date prévue.</p> <p>2. Accomplir les tâches d'évaluation : fixer des objectifs, des standards et des critères d'évaluation; choisir et élaborer des instruments d'évaluation des apprentissages des étudiants.</p> <p>2.1 Attribuer une pondération suffisamment importante aux activités d'évaluation pour être significative de la maîtrise globale des éléments de compétences, de la ou des compétences visées par le cours.</p> <p>2.2 Répartir l'évaluation sommative sur plusieurs activités d'évaluation en cours de session.</p>
Département	<p>1. S'assurer de la concertation des enseignants sur les objectifs, les contenus, les modes d'évaluation, les exigences et les seuils de réussite des cours et des programmes qui les concernent.</p> <p>1.1 Rechercher l'équivalence des évaluations par la concertation entre les professeurs d'un même département et entre les départements au sujet des modes et des instruments d'évaluation et à propos du degré d'exigence et des seuils de réussite.</p> <p>2. Élaborer des règles d'évaluation propres au programme qui précisent comment la maîtrise de la compétence est évaluée pour obtenir la note de passage.</p> <p>3. Prescrire dans le plan cadre une pondération suffisamment importante des activités d'évaluation pour être significative de la maîtrise globale des éléments de compétences, de la ou des compétences visées par le cours.</p> <p>4. S'assurer que les activités et les instruments d'évaluation respectent les principes de la politique institutionnelle d'évaluation des apprentissages (PIEA), les modalités départementales d'évaluation des apprentissages et le plan cadre.</p> <p>5. Valider, administrer et évaluer l'épreuve synthèse de programme.</p> <p>6. Se renseigner sur les besoins de perfectionnement ou de soutien en matière d'évaluation des apprentissages et planifier lui-même des activités de perfectionnement.</p> <p>6.1 Communiquer les besoins de perfectionnement ou de soutien en matière d'évaluation des apprentissages à la direction des études.</p>

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A2.2 L'échantillon des observations est suffisamment grand pour contrôler l'erreur aléatoire

Service de formation continue	1. Offrir des services d'encadrement et de perfectionnement au personnel enseignant pour les soutenir dans leurs pratiques d'évaluation des apprentissages.
Comité de programme	1. Assumer la responsabilité de l'élaboration de l'épreuve synthèse de programme et participer à son administration.
Direction des études	1. Lors de l'attribution d'une dispense, d'une équivalence ou d'une substitution, élaborer et administrer un examen démontrant la maîtrise par l'étudiant de la compétence et des éléments de compétence prévus au cours faisant l'objet de la demande. 2. S'assurer que soient organisées des activités de perfectionnement en matière d'évaluation des apprentissages. 3. Approuver les modalités générales de l'épreuve synthèse de chaque programme conduisant au diplôme d'études collégiales.
Commission des études	1. Donner son avis à la direction des études sur les modalités générales de l'épreuve synthèse de programme et sur l'harmonisation des pratiques d'évaluation dans l'ensemble des programmes.

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A3.1 Le score univers est lié au score cible

A3.1 Le score univers est lié au score cible.	
Professeur	<p>1. S'impliquer et participer et aux activités de concertation de programme.</p> <p>2. Planifier les activités d'apprentissage et d'évaluation en conformité avec la politique institutionnelle d'évaluation des apprentissages (PIEA), les règles départementales d'évaluation des apprentissages, le programme et le plan de cours cadre.</p> <p>2.1 Rédiger un plan de cours conforme à la politique institutionnelle d'évaluation des apprentissages (PIEA), aux règles départementales d'évaluation des apprentissages, au programme et au plan de cours cadre</p> <p>2.1.1 Indiquer dans le plan de cours la place du cours dans le programme d'études.</p> <p>2.1.2 Inclure au plan de cours une bibliographie et des lectures complémentaires aux cours qu'il enseigne.</p> <p>3. Accomplir les tâches d'évaluation : fixer des objectifs, des standards et des critères d'évaluation; choisir et élaborer des instruments d'évaluation des apprentissages des étudiants.</p> <p>3.1 Centrer l'évaluation sur les objectifs du programme.</p> <p>4. Justifier de la note accordée à l'évaluation par une rétroaction complète et détaillée.</p>
Département	<p>1. S'assurer de la concertation des enseignants sur les objectifs, les contenus, les modes d'évaluation, les exigences et les seuils de réussite des cours et des programmes qui les concernent.</p> <p>2. Participer à l'élaboration des objectifs du programme.</p> <p>2.1 S'assurer que les cours sont intégrés au programme et qu'ils contribuent à l'atteinte des objectifs du programme.</p> <p>3. Élaborer un plan de cours cadre pour chacun des cours dont il est responsable. <i>Le plan cadre présente les liens entre ce cours et les autres cours du programme.</i></p> <p>4. Pondérer l'évaluation des apprentissages en fonction de la valeur relative accordée aux objectifs d'apprentissage et aux compétences dans chaque cours dont il est responsable.</p> <ul style="list-style-type: none"> ▪ Exemple : Déterminer un pourcentage maximal de la note totale que peut valoir une activité d'évaluation. ▪ Exemple : Sauf exception, la réussite d'un cours ne peut reposer uniquement sur la somme de réussites fragmentaires. Chaque cours doit comporter une activité d'évaluation synthèse témoignant de l'atteinte de la ou des compétences. <p>5. Valider, administrer et évaluer l'épreuve synthèse de programme.</p>

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A3.1 Le score univers est lié au score cible

Département	6. Examiner les variations dans la distribution des scores des cours de même nature, d'un même cours dispensé par plusieurs professeurs ou des sections d'un même cours.
Service de formation continue	1. Guider et assister les nouveaux professeurs dans la préparation et l'organisation des cours en accord avec les objectifs, les standards et les politiques du programme et du département.
Comité de programme	<p>1. Développer les critères et les standards d'évaluation pour le programme.</p> <p>1.1 Déterminer les critères spécifiques aux compétences professionnelles du domaine.</p> <p>2. Recommander aux départements et à la Direction des études les mesures favorisant l'harmonisation des pratiques et l'équivalence des évaluations pour tous les cours du programme.</p> <p>3. Assurer la concordance entre les éléments de la politique institutionnelle d'évaluation des apprentissages (PIEA), le développement de programme et les plans de cours.</p> <p>4. Assumer la responsabilité de l'élaboration de l'épreuve synthèse de programme et participer à son administration.</p> <ul style="list-style-type: none"> ▪ Exemple : En lien avec le profil de sortie et la formation fondamentale, l'épreuve synthèse porte sur les compétences, les objectifs et les standards du programme d'études et non sur le contenu spécifique de chacun des cours du programme. ▪ Exemple : L'épreuve synthèse tient compte des objectifs et des standards déterminés dans le programme ministériel, du profil de sortie et de la perspective d'équivalence intra-institutionnelle de l'évaluation et de la synthèse.
Direction des études	<p>1. Réviser la politique institutionnelle d'évaluation des apprentissages (PIEA)</p> <p>1.1 Préciser dans la politique institutionnelle d'évaluation des apprentissages (PIEA) les éléments de formation qui doivent apparaître dans le plan de cours et faire l'objet d'évaluation.</p> <p>2. Élaborer la description du programme d'étude.</p> <p>2.1 Établir la liste des activités d'apprentissage pour chacun des programmes.</p> <p>3. Approuver les modalités générales de l'épreuve synthèse de chaque programme conduisant au diplôme d'études collégiales.</p> <p>4. Assurer l'administration de toute épreuve uniforme imposée par le ministre.</p> <p>5. Lors de l'attribution d'une dispense, d'une équivalence ou d'une substitution, élaborer et administrer un examen démontrant la maîtrise par l'étudiant de la compétence et des éléments de compétence prévus au cours faisant l'objet de la demande.</p> <p>6. Décrire les procédures par lesquelles les départements ou les programmes examinent les variations dans la distribution des scores des cours de même nature, d'un même cours dispensé par plusieurs professeurs ou des sections d'un même cours.</p>

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A3.1 Le score univers est lié au score cible

Registraire	1. Avant de recommander la graduation d'un étudiant au Ministère, vérifier si toutes les conditions de validation de l'argument interprétatif établies par le Ministère et par le conseil d'administration du collège ont été rencontrées.
Commission des études	1. Recommander à la direction des études les modifications qu'elle juge nécessaires à la politique institutionnelle d'évaluation des apprentissages (PIEA). 2. Donner son avis à la direction des études sur les modalités générales de l'épreuve synthèse de programme et sur l'harmonisation des pratiques d'évaluation dans l'ensemble des programmes.

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A3.2 Il n'y a pas d'erreur systématique qui soit susceptible de miner l'extrapolation

A3.2 Il n'y a pas d'erreur systématique qui soit susceptible de miner l'extrapolation.	
Étudiant	1. Chercher à s'assurer que la performance ne soit pas biaisée par des compétences auxiliaires. <ul style="list-style-type: none"> ▪ Exemple : S'assurer de posséder toute l'information nécessaire concernant les exigences linguistiques selon lesquelles il sera évalué. ▪ Exemple : Développer, durant ses études collégiales, les compétences linguistiques requises à son niveau d'études.
Professeur	1. Fournir aux étudiants les consignes de présentation des travaux. 2. Détecter la source d'erreur systématique et référer l'étudiant éprouvant des difficultés vers une ressource d'aide appropriée. <ul style="list-style-type: none"> ▪ Exemple : référer l'étudiant qui éprouve des difficultés en langue parlée et écrite au service d'aide en français du collège. 3. Adapter les activités d'évaluation aux besoins particuliers de certains groupes ou certains individus pour lesquels les activités d'évaluations planifiés pourraient constituer une source de biais systématique.
Département	1. S'assurer de la concertation des enseignants sur les objectifs, les contenus, les modes d'évaluation, les exigences et les seuils de réussite des cours et des programmes qui les concernent. <ul style="list-style-type: none"> 1.1 Rechercher l'équivalence des évaluations par la concertation entre les professeurs d'un même département et entre les départements au sujet des modes et des instruments d'évaluation et à propos du degré d'exigence et des seuils de réussite. 2. Déterminer les modalités spécifiques de présentation des travaux et les exigences propres aux laboratoires et aux stages. 3. Valider, administrer et évaluer l'épreuve synthèse de programme.
Comité de programme	1. Assumer la responsabilité de l'élaboration de l'épreuve synthèse de programme et participer à son administration.
Direction des études	1. Réviser la politique institutionnelle d'évaluation des apprentissages (PIEA) <ul style="list-style-type: none"> 1.1 Préciser dans la politique institutionnelle d'évaluation des apprentissages (PIEA) les éléments de formation qui doivent apparaître dans le plan de cours et faire l'objet d'évaluation. 2. Consulter la commission des études au sujet des règles d'évaluation propres aux programmes et de leur approbation. 3. Approuver les modalités générales de l'épreuve synthèse de chaque programme conduisant au diplôme d'études collégiales.

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A3.2 Il n'y a pas d'erreur systématique qui soit susceptible de miner l'extrapolation

Direction des études	4. Lors de l'attribution d'une dispense, d'une équivalence ou d'une substitution, élaborer et administrer un examen démontrant la maîtrise par l'étudiant de la compétence et des éléments de compétence prévus au cours faisant l'objet de la demande.
	5. Assurer l'administration de toute épreuve uniforme imposée par le ministre.
Commission des études	1. Recommander à la direction des études les modifications qu'elle juge nécessaires à la politique institutionnelle d'évaluation des apprentissages (PIEA).
	2. Donner son avis à la direction des études sur les modalités générales de l'épreuve synthèse de programme et sur l'harmonisation des pratiques d'évaluation dans l'ensemble des programmes.

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A4.1 Les implications associées au trait sont appropriées

A4.1 Les implications associées au trait sont appropriées.	
Professeur	<p>1. Rédiger un plan de cours conforme à la politique institutionnelle d'évaluation des apprentissages (PIEA), aux règles départementales d'évaluation des apprentissages, au programme et au plan de cours cadre.</p> <p>1.1 Préciser dans le plan de cours les exceptions (conditions particulières pouvant affecter le résultat qui ne sont pas nécessairement reliées à la compétence dans le domaine cible).</p> <ul style="list-style-type: none"> ▪ <i>Exemple : un objectif d'apprentissage a préséance sur les autres objectifs d'apprentissage en raison de son importance par rapport au domaine cible et s'il n'est pas maîtrisé, il peut à lui seul entraîner un échec.</i> ▪ <i>Exemple : intégrer aux objectifs de chacun de ses cours des objectifs d'ordre linguistique : maîtrise du vocabulaire, compréhension de textes écrits, communication verbale et écrite.</i> ▪ <i>Exemple : indiquer dans le plan de cours les modalités de participation au cours et les conséquences de l'absence aux activités d'apprentissage et d'évaluation sur la notation.</i> ▪ <i>Exemple : indiquer dans le plan de cours les conséquences du plagiat et de la tricherie sur la notation.</i> <p>1.2 Indiquer dans le plan de cours la pondération des activités d'évaluation.</p>
Département	<p>2. Justifier de la note accordée à l'évaluation par une rétroaction complète et détaillée.</p> <p>1. S'assurer de la concertation des professeurs sur les compétences auxiliaires et sur les implications associées à ces compétences.</p> <p>2. Préciser dans les règles départementales les exceptions aux conditions particulières pouvant affecter le résultat qui ne sont pas nécessairement reliées à la compétence dans le domaine cible.</p> <ul style="list-style-type: none"> ▪ <i>Exemple : Spécifier les cours ou les types de travaux ou d'examens dans lesquels l'évaluation du français écrit est inapplicable.</i> <p>3. Pondérer l'évaluation des apprentissages en fonction de la valeur relative accordée aux objectifs d'apprentissage et aux compétences dans chaque cours dont il est responsable.</p> <ul style="list-style-type: none"> ▪ <i>Exemple : Sauf exception, la réussite d'un cours ne peut reposer uniquement sur la somme de réussites fragmentaires. Chaque cours doit comporter une activité d'évaluation synthèse témoignant de l'atteinte de la ou des compétences.</i> ▪ <i>Exemple : Déterminer un pourcentage maximal de la note totale que peut valoir une activité d'évaluation.</i> <p>4. Valider, administrer et évaluer l'épreuve synthèse de programme.</p>

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A4.1 Les implications associées au trait sont appropriées

Comité de programme	<p>1. Assurer la définition du profil de sortie de l'étudiant inscrit dans ce programme.</p> <p>2. Assumer la responsabilité de l'élaboration de l'épreuve synthèse de programme et participer à son administration.</p> <ul style="list-style-type: none"> Exemple : En lien avec le profil de sortie et la formation fondamentale, l'épreuve synthèse porte sur les compétences, les objectifs et les standards du programme d'études et non sur le contenu spécifique de chacun des cours du programme.
Direction des études	<p>1. Réviser la politique institutionnelle d'évaluation des apprentissages (PIEA)</p> <p>1.1 Préciser dans la politique institutionnelle d'évaluation des apprentissages (PIEA) les exceptions (conditions particulières pouvant affecter le résultat qui ne sont pas nécessairement reliées à la compétence dans le domaine cible)</p> <ul style="list-style-type: none"> Exemple : un objectif d'apprentissage a préséance sur les autres objectifs d'apprentissage en raison de son importance par rapport au domaine cible et s'il n'est pas maîtrisé, il peut à lui seul entraîner un échec. Exemple : la langue écrite peut être considérée comme une compétence auxiliaire qui doit nécessairement être maîtrisée pour que la compétence dans le domaine cible soit entièrement atteinte. Exemple : la présentation matérielle des écrits est considérée comme une compétence auxiliaire qui doit nécessairement être maîtrisée pour que la compétence dans le domaine cible puisse être évaluée. Exemple : l'absence aux activités d'apprentissage et d'évaluation peut entraîner le retrait du droit de l'étudiant d'être évalué et conséquemment l'échec éventuel, et ce, quel que soit le niveau de compétence de l'étudiant dans le domaine cible. Exemple : la remise d'un travail incomplet ou le retard dans la remise d'un travail peuvent entraîner l'échec, et ce, quel que soit le niveau de compétence de l'étudiant dans le domaine cible. Exemple : le plagiat et la tricherie peuvent entraîner l'échec, et ce, quel que soit le niveau de compétence de l'étudiant dans le domaine cible. <p>2. Élaborer la description du programme d'étude.</p> <p>2.1 Établir la liste des activités d'apprentissage pour chacun des programmes.</p> <p>3. Approuver les modalités générales de l'épreuve synthèse de chaque programme conduisant au diplôme d'études collégiales.</p>

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A4.1 Les implications associées au trait sont appropriées

Direction des études	<p>4. Assurer l'administration de toute épreuve uniforme imposée par le ministre.</p> <p>5. Lors de l'attribution d'une dispense, d'une équivalence ou d'une substitution, élaborer et administrer un examen démontrant la maîtrise par l'étudiant de la compétence et des éléments de compétence prévus au cours faisant l'objet de la demande.</p> <ul style="list-style-type: none"> ▪ <i>Exemple : pour les cours touchés par une épreuve uniforme du ministre, la direction des études ne peut accorder une dispense, une équivalence ou une substitution que dans la mesure où l'étudiant réussit un examen qui démontre sa maîtrise de la compétence et des éléments de compétence prévus à ce cours.</i>
Commission des études	<p>1. Recommander à la direction des études les modifications qu'elle juge nécessaires à la politique institutionnelle d'évaluation des apprentissages (PIEA).</p>
Conseil d'administration	<p>1. Établir une liste des conditions devant être rencontrées pour valider l'argument interprétatif.</p> <ul style="list-style-type: none"> ▪ <i>Exemple : l'étudiant a réussi tous les cours ou tous les objectifs du programme.</i> ▪ <i>Exemple : l'étudiant a réussi l'épreuve synthèse de programme.</i> ▪ <i>Exemple : l'étudiant a réussi, le cas échéant, les épreuves imposées par le ministre dans des disciplines de la formation générale (ex. : Épreuve uniforme de français, English Exit Examination).</i> <p>1.1 Recommander au ministre de décerner un diplôme d'études collégiales à un étudiant.</p>

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A4.2 Les propriétés des scores observés supportent les implications associées avec l'étiquette du trait

A4.2 Les propriétés des scores observés supportent les implications associées avec l'étiquette du trait.	
Professeur	1. Justifier de la note accordée à l'évaluation par une rétroaction complète et détaillée.
Département	<p>1. Établir les modalités d'évaluation spécifiques et préciser les standards de réussite pour chacun des cours dont il est responsable.</p> <ul style="list-style-type: none"> ▪ <i>Exemple : La note de passage est de 60 %. Cette note traduit l'atteinte minimale des objectifs et des standards visés par le cours.</i> <p>1.1 Choisir entre une pondération plus importante de l'évaluation finale ou une répartition équivalente de la pondération des différentes activités d'évaluation ?</p> <ul style="list-style-type: none"> ▪ <i>Exemple I : Chaque cours comprend une évaluation finale dont la pondération est d'au moins 30 points de la note totale du cours. Si la pondération de l'évaluation finale est inférieure à 30 points de la note totale du cours, l'étudiant doit obtenir la note de passage à cette évaluation finale pour réussir le cours.</i> ▪ <i>Exemple II : la pondération de l'examen final ne doit pas excéder 25% de la note totale d'un cours.</i>
	<p>2. Pondérer l'évaluation des apprentissages en fonction de la valeur relative accordée aux objectifs d'apprentissage et aux compétences dans chaque cours dont il est responsable.</p> <ul style="list-style-type: none"> ▪ <i>Exemple : Sauf exception, la réussite d'un cours ne peut reposer uniquement sur la somme de réussites fragmentaires. Chaque cours doit comporter une activité d'évaluation synthèse témoignant de l'atteinte de la ou des compétences.</i> ▪ <i>Exemple : Déterminer un pourcentage maximal de la note totale que peut valoir une activité d'évaluation.</i>
	3. Concevoir, valider, administrer et évaluer l'épreuve synthèse de programme.
Comité de programme	<p>1. Assumer la responsabilité de l'élaboration de l'épreuve synthèse de programme et participer à son administration.</p> <ul style="list-style-type: none"> ▪ <i>Exemple : En lien avec le profil de sortie et la formation fondamentale, l'épreuve synthèse porte sur les compétences, les objectifs et les standards du programme d'études et non sur le contenu spécifique de chacun des cours du programme.</i>

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A4.2 Les propriétés des scores observés supportent les implications associées avec l'étiquette du trait

Direction des études	<p>1. Réviser la politique institutionnelle d'évaluation des apprentissages (PIEA)</p> <p>1.1 Préciser dans la politique institutionnelle d'évaluation des apprentissages (PIEA) les scores et mentions pouvant figurer au bulletin d'études collégiales.</p> <ul style="list-style-type: none"> ▪ Exemple : Après la passation de l'épreuve synthèse, la mention "réussite" ou "échec" paraîtra au bulletin d'études collégiales de l'étudiant. ▪ Exemple : le score associé à une réussite étant " RE ", les propriétés du score supportent les implications liés au fait que l'étudiant satisfait aux exigences de l'épreuve uniforme de français ou pour l'épreuve synthèse de programme. ▪ Exemple : le score associé à l'échec étant " EC ", les propriétés du score démontrent que l'étudiant a terminé un cours avec une note inférieure à la note de passage. ▪ Exemple : le score associé à l'équivalence étant " EQ ", les propriétés du score démontrent qu'il s'agit d'une équivalence et donc supportent les implications liées au fait qu'il s'agit d'une équivalence. ▪ Exemple : le score associé à un cours incomplet étant " IN ", les propriétés du score supportent les implications liées au fait qu'il s'agit d'un cours non complété. ▪ Exemple : le score associé à la substitution étant " SU ", les propriétés du score démontrent qu'il s'agit d'une substitution et donc supportent les implications liées au fait qu'il s'agit d'une substitution. ▪ Exemple : le score associé à un objectif atteint étant " OA ", les propriétés du score démontrent qu'il s'agit d'un objectif atteint et donc supportent les implications liées au fait que l'objectif a été atteint. ▪ Exemple : le score associé à un objectif techniquement atteint étant " OT ", les propriétés du score démontrent qu'il s'agit d'un objectif atteint à la suite d'une substitution, d'une équivalence ou d'une dispense et donc supportent les implications liées au fait que l'objectif a été atteint. ▪ Suggestion : accorder une mention particulière au relevé de note en cas d'abandon, par exemple "EA=échec par abandon".
----------------------	---

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A4.2 Les propriétés des scores observés supportent les implications associées avec l'étiquette du trait

Direction des études	<p>1.2 Préciser dans la politique institutionnelle d'évaluation des apprentissages (PIEA) les exceptions à la notation.</p> <ul style="list-style-type: none"> ▪ <i>Exemple : Aucun point ne peut être attribué pour la seule présence au cours, ni être retiré pour la seule absence aux cours.</i> ▪ <i>Exemple : Une pondération maximale de 10% peut être attribuée à la qualité de la langue. Dans les cours où la correction linguistique fait partie des objectifs d'apprentissage, la pondération pour la maîtrise de la langue écrite et parlée peut être supérieure à 10%.</i> <p>2. Approuver les modalités générales de l'épreuve synthèse de chaque programme conduisant au diplôme d'études collégiales.</p> <p>3. Assurer l'administration de toute épreuve uniforme imposée par le ministre.</p> <p>4. Lors de l'attribution d'une dispense, d'une équivalence ou d'une substitution, élaborer et administrer un examen démontrant la maîtrise par l'étudiant de la compétence et des éléments de compétence prévus au cours faisant l'objet de la demande.</p> <ul style="list-style-type: none"> ▪ <i>Exemple : pour les cours touchés par une épreuve uniforme du ministre, la direction des études ne peut accorder une dispense, une équivalence ou une substitution que dans la mesure où l'étudiant réussit un examen qui démontre sa maîtrise de la compétence et des éléments de compétence prévus à ce cours.</i>
----------------------	--

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages

A5.1 L'ensemble des arguments est respecté de façon à ce que chacune des inférences considérées de façon individuelle soit convaincante

A5.1 L'ensemble des arguments est respecté de façon à ce que chacune des inférences considérées de façon individuelle soit convaincante.	
Département	1. S'assurer que les pratiques d'évaluation sont conformes à la politique institutionnelle d'évaluation des apprentissages (PIEA).
Comité de programme	1. Lorsque le comité de programme est avisé du non respect de la politique institutionnelle d'évaluation des apprentissages (PIEA) par un étudiant ou un membre du personnel, prendre les mesures nécessaires pour découvrir les faits et, si nécessaire, corriger la situation.
Direction des études	1. Établir un portrait statistique de l'évaluation des apprentissages pour chaque département ou programme et de l'application de chaque argument de validité à la fin de chaque année.
Commission des études	1. Vérifier que les modalités départementales d'évaluation des apprentissages sont conformes à la politique institutionnelle d'évaluation des apprentissages (PIEA).

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A5.2 Les processus d'apprentissage et d'évaluation sont connus des acteurs

A5.2 Les processus d'apprentissage et d'évaluation sont connus des acteurs.	
Étudiant	1. Prendre connaissance de la politique institutionnelle d'évaluation des apprentissages (PIEA) et des modalités et des règles d'évaluation propres au cours et au programme d'études. ▪ <i>Exemple : S'assurer de posséder toute l'information nécessaire concernant les exigences linguistiques selon lesquelles il sera évalué.</i>
	2. Prendre connaissance du plan de cours.
	3. Prendre connaissance du calendrier des évaluations pour l'ensemble du collège mis en place par la direction des études.
	4. Participer aux activités d'apprentissage en classe et, à défaut d'être présent, de récupérer par lui-même les apprentissages manqués.
	5. Assurer un suivi personnel de ses apprentissages, de ses résultats et de son cheminement scolaire.
	5.1 Ajuster sa démarche d'apprentissage en utilisant les rétroactions obtenues et les résultats de ses évaluations. 5.2 Au besoin, rechercher de l'aide auprès des services et du personnel du collège.
Professeur	1. Planifier les activités d'apprentissage et d'évaluation de façon à préparer progressivement l'étudiant à l'épreuve synthèse de programme et à l'atteinte des objectifs du cours et du programme.
	2. Informer l'étudiant sur le programme, sur les activités d'apprentissage et sur les modes d'évaluation des apprentissages.
	2.1 Rédiger un plan de cours conforme à la politique institutionnelle d'évaluation des apprentissages (PIEA), aux règles départementales d'évaluation des apprentissages, au programme et au plan de cours cadre.
	2.1.1 Indiquer dans le plan de cours les standards de performance requis pour réussir le cours. 2.1.2 Indiquer dans le plan de cours les principales règles d'évaluation des apprentissages ainsi que les moyens d'accès à l'information complète concernant la politique institutionnelle d'évaluation des apprentissages (PIEA) et les modalités départementales d'évaluation des apprentissages. ▪ <i>Exemple : indiquer dans le plan de cours les conséquences du plagiat et de la tricherie sur la notation.</i>

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A5.2 Les processus d'apprentissage et d'évaluation sont connus des acteurs

Professeur	<p>2.1.2.1 Indiquer dans le plan de cours les modalités de participation au cours et les conséquences de l'absence aux activités d'apprentissage et d'évaluation sur la notation.</p> <p>2.1.3 Indiquer dans le plan de cours les méthodes pédagogiques utilisées et le déroulement des activités.</p> <p>2.1.4 Indiquer dans le plan de cours les activités d'évaluation sommative des apprentissages et leurs modalités : la nature des épreuves (travaux, examens ou autres) servant à l'établissement de la note finale, leur pondération, leur durée et la date prévue.</p> <p>2.1.4.1 Indiquer dans le plan de cours les dates de remise des travaux, des rapports de laboratoires et des examens ainsi que le nombre d'activités d'évaluation des apprentissages.</p> <ul style="list-style-type: none"> ▪ <i>Exemple : aviser les étudiants au moins deux semaines à l'avance de la tenue de toute évaluation dont la pondération excède 5% de la note globale du cours.</i> <p>2.1.5 Remettre le plan de cours à l'étudiant au début du cours.</p> <p>2.1.5.1 Négocier avec les étudiants tout changement au plan de cours en cours de session et les en informer par écrit.</p> <p>3. S'assurer que l'étudiant connaisse et comprenne les conditions menant à la réussite ou à l'échec du cours et du programme</p> <ul style="list-style-type: none"> ▪ <i>Exemple : s'assurer que l'étudiant comprenne ce que constitue le plagiat dans le cadre du cours et les conséquences de cette pratique.</i> ▪ <i>Exemple : indiquer dans le plan de cours les conséquences du plagiat et de la tricherie sur la notation.</i> <p>3.1 Indiquer dans le plan de cours les standards de performance requis pour réussir le cours.</p> <p>3.2 Fournir à l'avance les critères d'évaluation à l'étudiant.</p> <p>4. Informer l'étudiant sur son processus d'apprentissage et sur le développement et l'acquisition des compétences propres à son programme.</p> <ul style="list-style-type: none"> ▪ <i>Exemple : au moyen de l'évaluation formative, de notes partielles, d'auto-évaluations ou de coévaluations, etc.</i> <p>4.1 Informer l'étudiant des processus d'évaluation employés et des résultats de l'évaluation.</p>
------------	--

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.

Moyens pour chercher à assurer la validité des inférences d'évaluation des apprentissages
A5.2 Les processus d'apprentissage et d'évaluation sont connus des acteurs

Professeur	<p>4.1.1 Communiquer à l'étudiant les résultats des activités d'évaluation dans un délai raisonnable.</p> <ul style="list-style-type: none"> Exemple : un maximum de deux semaines après l'évaluation pour les tests, les intras et les courts travaux et un maximum de trois semaines pour les travaux longs. <p>4.1.2 En cas d'échec, poser un diagnostic pour évaluer les lacunes de l'étudiant et ainsi assurer sa préparation à une prochaine évaluation ou à une reprise.</p> <p>4.1.3 Justifier de la note accordée à l'évaluation par une rétroaction complète et détaillée.</p> <p>4.2 Donner de façon continue de la rétroaction à l'étudiant sur l'évolution de ses apprentissages.</p> <p>4.2.1 Justifier de la note accordée à l'évaluation par une rétroaction complète et détaillée.</p> <p>4.2.2 Fournir à l'étudiant une estimation de ses chances de réussite.</p>
Département	<p>1. Élaborer un plan de cours cadre pour chacun des cours dont il est responsable.</p> <p>1.1 Le plan cadre présente les étapes de l'apprentissage et les approches pédagogiques suggérées.</p> <p>1.2 Le plan cadre présente les modalités d'évaluation incluant l'épreuve finale ou ce qui en tient lieu.</p> <p>2. Prévoir des procédures de gestion de conflits en lien avec la politique institutionnelle d'évaluation des apprentissages (PIEA).</p>
Direction des études	<p>1. Réviser la politique institutionnelle d'évaluation des apprentissages (PIEA)</p> <p>1.1 Préciser dans la politique institutionnelle d'évaluation des apprentissages (PIEA) les droits et responsabilités de tous les intervenants dans le processus d'évaluation des apprentissages.</p> <p>2. Veiller à la diffusion de la politique institutionnelle d'évaluation des apprentissages (PIEA) auprès des étudiants, des professeurs et du personnel concerné ainsi qu'auprès des individus et organismes intéressés à l'extérieur du collège.</p> <p>3. Élaborer la description du programme d'étude.</p> <p>3.1 Fournir la description du programme d'étude et la grille des cours à l'étudiant.</p> <p>4. Informer l'étudiant sur le programme, sur les conditions à la graduation, sur les exigences des universités et du marché du travail.</p> <p>5. Mettre en place et diffuser un calendrier des évaluations pour l'ensemble du collège.</p> <p>6. Remettre à l'étudiant un bulletin de mi-session et un bulletin de fin de session.</p>
Conseil d'administration	<p>1. Attester auprès des élèves, du public, du Ministre et de la Commission d'évaluation de l'enseignement collégial de la qualité, de l'équité et de l'équivalence des évaluations des apprentissages est un moyen pour le conseil d'administration de chercher à s'assurer que les processus d'évaluation sont connus et acceptés des acteurs.</p>

Note : les éléments cernés en gris sont des éléments spécifiques à cet argument qui ne se répètent dans aucun autre argument de la structure d'argumentation de la validité des inférences d'évaluation.